



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

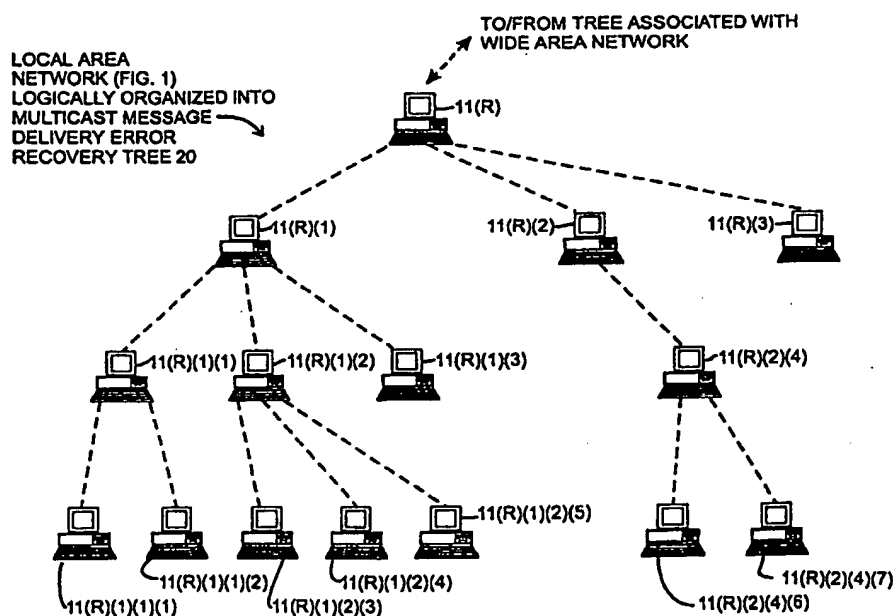
(51) International Patent Classification ⁶ : H04L 12/18	A1	(11) International Publication Number: WO 99/55042 (43) International Publication Date: 28 October 1999 (28.10.99)
(21) International Application Number: PCT/US99/07750 (22) International Filing Date: 8 April 1999 (08.04.99) (30) Priority Data: 09/061,849 18 April 1998 (18.04.98) US (71) Applicant: SUN MICROSYSTEMS, INC. [US/US]; 901 San Antonio Road, Palo Alto, CA 94303 (US). (72) Inventors: CHIU, Dah, Ming; 736 Main Street, Acton, MA 01720 (US). KADANSKY, Miriam; 165 Westview Drive, Westford, MA 01886 (US). PERLMAN, Radia, J.; 10 Huckleberry Lane, Acton, MA 01720 (US). (74) Agents: HOOVER, George, W. et al.; Blakely, Sokoloff, Taylor & Zafman, 7th floor, 12400 Wilshire Boulevard, Los Angeles, CA 90025-1026 (US).		(81) Designated States: AE, AL, AM, AT, AT (Utility model), AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, CZ (Utility model), DE, DE (Utility model), DK, DK (Utility model), EE, EE (Utility model), ES, FI, FI (Utility model), GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK (Utility model), SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>

(54) Title: SYSTEM AND METHOD FOR ESTABLISHING A MULTICAST MESSAGE DELIVERY ERROR RECOVERY TREE IN A DIGITAL NETWORK

(57) Abstract

In a digital data network, a plurality of devices interconnected by a communication link organize themselves into a tree structure. Each of the devices has an associated suitability value that generally relates to the device's suitability for becoming a node in the tree structure. The devices organize themselves into a tree structure in one or more iterations, each iteration comprising two general steps, namely, a node election step and a tree establishment step. In the node election step, the devices whose suitability values are such that they can become nodes in the tree broadcast over the communication link node election messages including their respective suitability values.

These devices also receive the node election messages that are broadcast by other devices. Each device determines whether it is elected a node in the tree structure in connection with a comparison between its suitability value and suitability values of node election messages received thereby. During the tree establishment step, the devices in the network communicate with at least one of the device or devices which is or are elected respective nodes in the tree structure to facilitate becoming respective children thereof.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

SYSTEM AND METHOD FOR ESTABLISHING A MULTICAST MESSAGE
DELIVERY ERROR RECOVERY TREE IN A DIGITAL NETWORK

-1-

1 FIELD OF THE INVENTION

2 The invention relates generally to the field of digital networks and more particularly provides
3 an efficient arrangement for establishing a multicast message delivery error recovery tree in a digital
4 network.

5 BACKGROUND OF THE INVENTION

6 In modern "enterprise" digital data processing systems for use in an office environment in
7 a company, a number of personal computers, workstations, and other various network resources such
8 as mass storage subsystems, network printers and interfaces to the public telephony system, are
9 typically interconnected in a computer network. The personal computers and workstations are used
10 by individual users to perform processing in connection with data and programs that may be stored
11 in the network mass storage subsystems. In such an arrangement, the personal
12 computers/workstations, operating as clients, download the information, including data and
13 programs, from the network mass storage subsystems for processing. In addition, the personal
14 computers or workstations will enable processed data to be uploaded to the network mass storage
15 subsystems for storage, to a network printer for printing, to the telephony interface for transmission
16 over the public telephony system, or the like. In such an arrangement, the network mass storage
17 subsystems, network printers and telephony interfaces operate as shared resources, since they are
18 available to service requests from all of the clients in the network. By organizing the network in
19 such a manner, the servers are readily available for use by all of the personal computers/workstations
20 in the network. Networks may be spread over a fairly wide area, and may interconnect personal
21 computers, workstations and other devices among a number of companies and individuals.

22 Computers and other resources (generally, "devices") in a network share information by
23 transferring messages thereamong. Messages can be categorized into two general classes, namely,

-2-

1 unicast messages and multicast messages. A unicast message is used by a device to transfer
2 information to one other device in the network, whereas a multicast messages is used by a device
3 to transfer information to all of the devices, or a selected subset of the devices, in the network. Each
4 unicast message includes address information, including a source address, which identifies the
5 particular device that transmitted the message, and a destination address, which identifies the
6 particular device that is to receive the message in the case of a unicast message. Similarly, each
7 multicast message includes address information, including a source address, which identified the
8 particular device that transmitted the message, and a destination address that identifies a set of
9 devices (which may be either all of the devices or a selected subset) that are to receive the message
10 in the case of a multicast message. When a particular device determines that a unicast message
11 contains a destination address that identifies it (that is, the particular device), or that a multicast
12 message contains a destination address that identifies a set of devices that includes the particular
13 device, it (that is, the particular device) will receive and process the message.

14 Although networks generally transfer messages reliably, sometimes messages are lost and
15 not received by the destination devices, that is, the particular devices that are intended to receive
16 them. If a destination device fails to receive a message, it can transmit a "negative acknowledgment"
17 ("NACK") message to the source device, that is, the device that generated and transmitted the
18 message over the network, requesting retransmission of the message. Typically, unicast messages
19 transmitted by a particular source device to a particular destination device include sequencing
20 information so that, if a destination device determines that there is a gap in the sequence of messages
21 received by it from a particular source device, it can notify the source device to request
22 retransmission. If messages can be delivered to the destination device out of the order in which they
23 are transmitted, the destination device may wait for a period of time after it notices the gap in the
24 sequence to see if the missing messages might be eventually delivered, but, in any case, if the
25 missing message are not delivered to the destination device at least prior to expiration of the period
26 of time, the destination device will generate the "NACK" message and transmit it to the source
27 device to enable retransmission of the missing messages.

-3-

1 A similar operation occurs in connection with multicast messages. That is, if a destination
2 device determines that a multicast message whose destination address identifies a set that includes
3 the destination device, in a sequence of such multicast messages from the same source device, has
4 not been received, it (that is the destination device) will generate a "NACK" message for
5 transmission to the source device requesting retransmission of the message, or at least transmission
6 of the information in the message in a unicast message to the destination device. However, unless
7 failure to receive a multicast message by one destination device is due to a malfunction of that
8 destination device, if one destination device fails to receive a multicast message, frequently a large
9 number or all of destination devices in the destination device set for the multicast message will also
10 fail to receive the multicast message, all of which will generate and transmit "NACK" messages to
11 the source device requesting retransmission of the multicast message. In that case, if there are a
12 large number of destination devices in the device set for the multicast message, the source device
13 may receive and process a large number of such "NACK" messages, which can significantly
14 negatively effect its (that is, the source device's) other processing operations.

15 SUMMARY OF THE INVENTION

16 The invention provides a new and improved arrangement for establishing a multicast
17 message delivery error recovery tree in a digital network which may be used to reduce the number
18 of "negative acknowledgment" ("NACK") messages which may be generated for transmission to a
19 particular source device, which generates a multicast message for transmission to a plurality of
20 destination devices, by destination devices if they fail to receive a multicast message generated and
21 transmitted by the source device.

22 In brief summary, the invention provides a system for, and a method of, logically organizing,
23 in a digital data network comprising a plurality of devices interconnected by a communication link,
24 into a tree structure. Each of the devices has an associated suitability value that generally relates to
25 the device's suitability for becoming a node in the tree structure. The devices organize themselves

-4-

1 into a tree structure in one or more iterations, each iteration comprising two general steps, namely,
2 a node election step and a tree establishment step. In the node election step, the devices whose
3 suitability values are such that they can become nodes in the tree broadcast over the communication
4 link node election messages including their respective suitability values. These devices also receive
5 the node election messages that are broadcast by other devices. Each device determines whether it
6 is elected a node in the tree structure in connection with a comparison between its suitability value
7 and suitability values node election messages received thereby. During the tree establishment step,
8 the devices in the network communicate with at least one of the device or devices which is or are
9 elected respective nodes in the tree structure to facilitate becoming respective children thereof.

10 BRIEF DESCRIPTION OF THE DRAWINGS

11 This invention is pointed out with particularity in the appended claims. The above and
12 further advantages of this invention may be better understood by referring to the following
13 description taken in conjunction with the accompanying drawings, in which:

14 FIG. 1 is a schematic diagram of a local area network including computers which establish
15 and organize a multicast message delivery error recovery tree in accordance with the invention;

16 FIG. 2 is a logical diagram of the local area network depicted in FIG. 1, logically organized
17 in an illustrative multicast message delivery error recovery tree;

18 FIGS. 3A and 3B depict structures of messages used in establishing a multicast message
19 delivery error recovery tree in connection with the invention; and

20 FIGS. 4A through 4G comprise flow diagrams depicting operations performed by the
21 computers in the local area network depicted in FIG. 1, in establishing an multicast message delivery
22 error recovery tree in accordance with the invention.

DETAILED DESCRIPTION OF AN ILLUSTRATIVE EMBODIMENT

FIG. 1 is a schematic diagram of a local area network 10 including an arrangement for generating or constructing a multicast tree in the network, in accordance with the invention. With reference to FIG. 1, local area network 10 includes a plurality of computers 11(1) through 11(N) (generally identified by reference numeral 11(n)) interconnected by a communication link 13. As is conventional, at least some of the computers 11(1), ..., 11(N-1) are in the form of personal computers or computer workstations, each of which includes a system unit, a video display unit and operator input devices such as a keyboard and mouse. The computer 11(N) also includes a system unit, and may also include a video display unit and operator input devices. The computers 11(n) are of the conventional stored-program computer architecture. A system unit generally includes processing, memory, mass storage devices such as disk and/or tape storage elements and other elements (not separately shown), including network interface devices represented by respective arrows 14(n) for interfacing the respective computer to the communication link 13. A video display unit permits the computer to display processed data and processing status to the user, and an operator input device enable the user to input data and control processing by the computer. The computers 11(n) transfer information, in the form of messages, through their respective network interface devices 14(n) among each other over the communication link 13.

The communication link 13 interconnecting the computers 11(n) in the network 10 may, as is conventional, comprise wires, optical fibers or other transmission media, and may further comprise switches, for, respectively, carrying and switching signals representing messages among the computers 11(n). As noted above, each of the computers 11(n) typically includes a network interface device 14(n), which connects the respective computer to the communications link 13. The transmission media and switches comprising communication link 13 may interconnect the computers 11 in any convenient topology.

-6-

1 Information is transferred among the computers 11(n) in the form of messages. Each
2 message contains a header portion, which generally contains information that is useful in controlling
3 the transfer of the message from the source computer 11(s), that is, the computer 11(n) that transmits
4 the message, to the destination computer 11(d), that is, the computer or computers that is/are to
5 receive the message, and a data portion, which generally contains information that is to be
6 transferred. The information contained in the header portion includes message transfer protocol
7 information, including, inter alia, source and destination addresses that identify the source computer
8 11(s) and the destination computer(s) 11(d) that is/are to receive the message, and each computer
9 11(n) can determine from a message's destination address whether it is to receive the message. The
10 destination address information in a message may identify one computer 11(n) which is to receive
11 the message, in which case the message is referred to as a "unicast" message. On the other hand, the
12 destination address information in a message may comprise a multicast address, which enables all
13 of the computers 11(n) in the local area network 10 to receive the message; a message which
14 includes such a multicast address is referred to as a "multicast" message or, alternatively, a
15 "broadcast" message. A multicast address may, alternatively, enable various subsets of the
16 computers 11(n) to receive the message; a message which contains such destination address
17 information will also be referred to as a multicast message. Multicast addresses provide a
18 mechanism by which a source computer 11(s) can transmit a single multicast message, and enable
19 a plurality of computers 11(n) to receive the message as destination computers.

20 The message transfer protocol information that is provided in the header of a message can
21 also include so-called "time to live" ("TTL") information, which can limit the scope of transmission
22 of a message. The time to live information may be used to limit the scope of transmission of a
23 message to, for example, a particular local area network, such as local area network 10, to various
24 groupings of contiguous local area networks in the wide area network, or to the entire wide area
25 network. If, for example, a computer 11(n) in the local area network 10 transmits a message that
26 contains a multicast address and time to live information indicating that the scope of transmission
27 is limited to the local area network 10, then the message will be received and processed only by the

-7-

1 computers 11(n) in the local area network 10 which are conditioned to respond to that multicast
2 address. Typically, the time to live information is used to control whether gateways (not shown)
3 which are used to interconnect local area networks will transmit a message that they receive on one
4 local area network onto another local area network. In that case, the time to live information is in
5 the form of an integer numerical value, and, when a gateway receives a message, it will decrement
6 the time to live value and, if the decremented time to live value is greater than zero, transmit the
7 message onto the other local area network. On the other hand, if the decremented time to live value
8 is zero, the gateway will not transmit the message received from the one local area network onto the
9 other local area network. Thus, computer 11(n) can limit the scope of transmission of a message to
10 local area network by transmitting a time to live value of "one," in the message.

11 The local area network 10 schematically depicted in FIG. 1 comprises a part of a wide area
12 network, such as a private wide area network maintained by a particular enterprise and/or a public
13 wide area network such as the Internet. Similar to the local area network 10, the other portions of
14 the wide area network that are not depicted in FIG. 1 generally include local area networks, each
15 including computers interconnected by respective communication links, and the various local area
16 networks are interconnected by various combinations of communication links and switches in a
17 conventional manner. The computers within each local area network can share information by
18 transferring respective unicast or multicast messages in a manner similar to that described above in
19 connection with local area network 10. In addition, a computer in one local area network can share
20 information with computers in other local area networks by transmitting respective unicast or
21 multicast messages in a manner similar to that described above in connection with local area network
22 10.

23 The invention provides an arrangement for facilitating the logical organization of at least
24 some of the computers 11(n) in the network 10 into a multicast message delivery error recovery tree.
25 Generally, a multicast message delivery error recovery tree is used to provide for efficient error
26 recovery in the event that one or more computers in the network (which may comprise either the
27 local area network 10 or a wide area network including local area network 10) fail to receive a

-8-

1 multicast message transmitted by a source computer 11(s). Typically, if a computer 11(n) is to be
2 a destination computer 11(d) to receive a message that is transmitted by a source computer 11(s),
3 and if the destination computer 11(d) fails to receive the message, it will generate a "negative
4 acknowledgment" ("NACK") message for transmission to the source computer 11(s) to enable the
5 source computer 11(s) to retransmit the message to the destination computer 11(d). This may be
6 repeated several times until the destination computer 11(d) actually receives the message. A
7 destination computer 11(d) can, for example, use message sequencing information that is generally
8 provided in the protocol information in the headers of respective messages in a sequence transmitted
9 by the source computer 11(s) to determine whether it failed to receive a particular message.

10 The use of "NACK" messages by a destination computer 11(d) to request retransmission of
11 messages which it did not receive is generally efficient in connection with unicast messages, since
12 there is only one destination computer 11(d) which may generate a "NACK" message if a message
13 is not received. In that case, the source computer 11(s) will only need to process one "NACK"
14 message for each non-received message. A problem can arise, however, if a number of destination
15 computers 11(d) do not receive a multicast message that is transmitted by a source computer 11(s),
16 and generate respective "NACK" messages for transmission to the source computer 11(s). In that
17 case, the source computer 11(s) would need to process each of the "NACK" messages that it
18 receives, which can significantly negatively affect other processing operations that are to be
19 performed by the source computer 11(s). A multicast message delivery error recovery tree can assist
20 in reducing the number of "NACK" messages which are transmitted to the source computer 11(s)
21 of a multicast message, which can serve to reduce the negative affect on processing thereby which
22 might otherwise arise, particularly in a wide-area network in which large numbers of computers may
23 be destinations of a multicast message. The invention provides an arrangement for efficiently
24 organizing the computers 11(n) in the local area network into a multicast message delivery error
25 recovery tree.

26 Before proceeding further, it would be helpful to describe the logical organization of a
27 multicast message delivery error recovery tree in local area network 10. An illustrative example of

-9-

1 a multicast message delivery error recovery tree, identified by reference numeral 20, is depicted in
2 FIG. 2. Generally, in the illustrative multicast message delivery error recovery tree 20, the
3 computers 11(n) comprising local area network 10 are "logically interconnected" (as will be
4 described below) in a manner represented by the dashed lines interconnecting the various computers
5 depicted in FIG. 2. In the illustrative multicast message delivery error recovery tree 20:

6 (i) one of the computers 11(R) forms a root node, which forms the highest level in the tree
7 20,

8 (ii) others of the computers 11(R)(1) through 11(R)(A) (generally identified by reference
9 numeral 11(R)(a), "a" being an index between "1" and "A," inclusive) all of which are "logically
10 connected" (as will be described below) to the root node 11(R), form a first lower level in the tree
11 20 below the root node,

12 (iii) others of the computers 11(R)(a)(1) through 11(R)(a)(B) (generally identified by
13 reference numeral 11(R)(a)(b), "b" being an index between "1" and "B," inclusive), all of which are
14 logically connected to a computer 11(R)(a) in the first level, form a second lower level in the tree
15 20 below the root node,

16 and so on, until finally the rest of the computers 11(R)(a)(b)...(1) through 11(R)(a)(b)...(L) (generally
17 identified by reference numeral 11(R)(a)(b)...(l), "l" being an index between "1" and "L," inclusive)
18 form respective "leaf nodes" in the tree 20. (It will be appreciated that the computers 11(R),
19 11(R)(a), 11(R)(a)(b), etc, comprise respective ones of computers 11(n) depicted in FIG. 1, the
20 indices "(n)" used in the reference numerals in FIG. 1 being changed in FIG. 2 to depict the
21 organization of the computers 11(n) in FIG. 1 into the multicast message delivery error recovery tree
22 20.) Generally,

23 (a) a computer that is in a level in the tree below the root node, or that forms a leaf node is
24 logically connected to one computer in a higher level, which computer will be referred to as its
25 "parent" in the multicast message delivery error recovery tree 20, and

-10-

1 (b) a computer that is in a level that is above the lowest level in the tree is logically connected
2 to one or more computers in the next lower level and/or to one or more computers that form
3 respective leaf node(s) in the multicast message delivery error recovery tree 20, which will be
4 referred to as the computer's "children."

5 Thus, with reference to the illustrative multicast message delivery error recovery tree 20 depicted
6 in FIG. 2, computer 11(R), which forms the root node in the tree, is the parent of the computers
7 11(R)(1) and 11(R)(2), which form the second level in the tree, and additionally is the parent of
8 computer 11(R)(3) which forms a leaf node in the tree 20. Contrariwise, each computer 11(R)(1),
9 11(R)(2) and 11(R)(3) is a "child" of the computer 11(R) in the tree 20. Similarly, computer
10 11(R)(1) is the parent of computers 11(R)(1)(1) and 11(R)(1)(2) (which comprise nodes in the third
11 level of tree 20) and 11(R)(1)(3) (which is a leaf node in tree 20), and each computer 11(R)(1)(1),
12 11(R)(1)(2) and 11(R)(1)(3) is a child of computer 11(R)(1). It will be appreciated that a computer
13 other than the computer that forms the root node will have one parent, but it can have one or more
14 than one children.

15 Generally, a computer that forms a leaf node in the multicast tree is logically connected to
16 a computer in a higher level, but no computer in any lower level and no computer forming a leaf
17 node logically connects to it (that is, the computer 11(R)(a)(b)...(l) that forms the leaf node). A
18 computer 11(R)(a)(b)...(l) that forms a leaf node may connect to a computer at any level in the tree.
19 It will be appreciated that the tree need not be balanced, that is, there may be different numbers of
20 levels between the computer 11(R) which forms the root node, and the various computers
21 11(R)(a)(b)...(l) which form leaf nodes in the tree, and various non-leaf nodes comprising the tree
22 20 may have different numbers of children. Thus, for example, the number of levels between the
23 computer 11(R) which forms the root node and a computer 11(R)(a)(b)...(l₁) which forms a leaf node,
24 may differ from the number of levels between the computer 11(R) and another computer
25 11(R)(a)(b)...(l₂) which forms a leaf node, and so forth for all computers 11(l) which form leaf nodes.
26 In each reference numeral 11(R)(a), 11(R)(a)(b), ...11(R)(a)(b)...(l), the sequence of values for
27 indices "a," "b," "l" define a logical path through the tree from the computer system 11(R) forming

-11-

1 the root node, through a sequence of computers in successively lower levels to the computer
2 identified by the respective reference numeral.

3 As noted above, the local area network 10 in which the multicast message delivery error
4 recovery tree 20 is formed is part of a wide area network including a plurality of local area networks.
5 In one embodiment, the multicast message delivery error recovery tree 20 depicted in FIG. 2 forms
6 a sub-tree of a larger tree (not shown) as represented by the double-headed arrow associated with the
7 legend "TO/FROM TREE ASSOCIATED WITH WIDE AREA NETWORK" in FIG. 2. In that
8 case, although the computer 11(R) which forms the root node in the multicast message delivery error
9 recovery tree 20 for local-area network 10 does not have a parent within the tree 20 organized as
10 depicted in FIG. 2, it will have a parent in the larger tree. The parent, and computers in levels
11 thereabove in the larger tree formed in the wide area network, may be in other local area networks
12 in the wide area network. Alternatively or in addition, computers at higher levels in the tree formed
13 in the wide area network may include computers in the local area network 10. In that case, a
14 computer in the local area network 10 may be at a node at a lower level or comprise a leaf node
15 11(R)(a)(b)...(l) in the multicast message delivery error recovery tree 20, and may in addition
16 comprise a node at a higher level than the computer 11(R) that forms the root node in the multicast
17 message delivery error recovery tree 20 formed in the local area network 10.

18 The multicast message delivery error recovery tree 20 is used to provide a mechanism
19 whereby, in the event that one or more of the computers in the network 10 fail to receive a multicast
20 message from the computer which forms the root node in the tree formed in the wide area network,
21 of which the local area network 10 is a part, those computers can efficiently request the re-
22 transmission of a copy of the multicast message thereto. For example, the tree 20 provides a
23 mechanism whereby a computer, such as computer 11(R)(1)(1)(1), in the network 10, if it fails to
24 receive a multicast message from another computer as source computer 11(s), instead of generating
25 a "NACK" message for transfer to the source computer 11(s) requesting re-transmission of a copy
26 of the multicast message, can generate a "NACK" message for transfer over communication link 13
27 to its parent computer 11(R)(1)(1) in multicast message delivery error recovery tree 20, the "NACK"

-12-

1 message requesting it (that is, the parent computer 11(R)(1)(1)) to provide a copy of the multicast
2 message to the child computer 11(R)(1)(1)(1). If that parent computer 11(R)(1)(1) has a copy of the
3 multicast message, it (that is, parent computer 11(R)(1)(1)) will transmit the multicast message over
4 the communication link 13. The child computer 11(R)(1)(1)(1) which requested the copy of the
5 multicast message can then receive the multicast message transmitted by the parent computer
6 11(R)(1)(1). It will be appreciated that the other computers 11(n) in the local area network 10 can
7 also receive the multicast message and use it if they had not previously received it; the other
8 computers 11(n) in the local area network, that is, the computers which previously received the
9 message, can ignore the message. Preferably, the parent computer 11(R)(1)(1), when it transmits
10 the multicast message, will transmit the message with a time to live value that will limit its
11 transmission to the local area network 10, so that the message is not transmitted to other portions of
12 the wide area network which may have received the multicast message.

13 On the other hand, if the computer 11(R)(1)(1) does not have a copy of the multicast
14 message, it will generate a "NACK" message for transfer over communication link 13 to its parent
15 computer 11(R)(1) requesting a copy thereof. If that computer 11(R)(1) has a copy of the
16 multicast message, it (that is, the computer 11(R)(1)) will transmit the multicast message over the
17 communication link 13 as described above, to provide the multicast message to its child and
18 grandchild computers 11(R)(1)(1) and 11(R)(1)(1)(1), along with any other computers 11(n) which
19 had not previously received the message. Similarly, if the computer 11(R)(1) does not have a copy
20 of the multicast message, it, in turn, will generate a "NACK" message for transfer over
21 communication link 13 to its parent, namely, the computer 11(R) which forms the root node in the
22 multicast tree 20, requesting a copy thereof. If the computer 11(R) has a copy of the requested
23 multicast message, it will transmit the multicast message over the communication link 13 as
24 described above.

25 However, if, after receiving the request from the computer 11(R)(1), the computer 11(R)
26 which forms the root node in the multicast message delivery error recovery tree 20 determines that
27 it does not have a copy of the requested message, it will attempt to obtain a copy. As noted above,

-13-

1 in one embodiment, in which the local area network 10 forms part of a wide area network, and in
2 which the multicast message delivery error recovery tree 20 forms a sub-tree of a larger tree formed
3 in the wide area network, if the computer 11(R) which forms the root node in the multicast message
4 delivery error recovery tree 20 does not have a copy of the message, it will generate a "NACK"
5 message for transfer to a computer in the wide area network that is its parent in the larger tree. This
6 process, in which each computer which receives such a "NACK" message from its child in the tree
7 determines whether it has a copy of the requested message, and (i) if so, transmits the multicast
8 message, but (ii) if not generates a "NACK" message for transfer to its parent, will continue until
9 a "NACK" message reaches the computer comprising root node in the tree formed in the wide area
10 network, which is the source computer 11(s) for the message.

11 It will be appreciated that the use of the multicast message delivery error recovery tree 20,
12 and associated tree formed in the wide area network, can serve to substantially reduce the number
13 of "NACK" messages that are transmitted to the source computer 11(s) for a multicast message. If
14 a multicast message is not received by a particular destination computer, it is frequently the case that
15 a large number of destination computers 11(d) will not receive the multicast message, not just a
16 single destination computer. In that case, the destination computers which fail to receive the
17 multicast message will generate only one "NACK" message. Thus, for example, if a computer
18 which forms a node in the multicast message delivery error recovery tree 20 has already generated
19 a "NACK" message requesting a copy of a particular multicast message for transmission to its parent
20 in the tree 20, either because it itself failed to receive the multicast message or in response to receipt
21 of a "NACK" message from one of its children requesting the multicast message, and it later receives
22 a "NACK" message from another child requesting the same multicast message, it will not need to
23 generate another "NACK" message for the same multicast message for transmission to its parent.
24 In addition, the "NACK" messages that are generated by the destination computers 11(d) are
25 generally not for transmission to the source computer 11(s) which generated the multicast message,
26 but instead to their respective parents in the multicast message delivery error recovery tree 20, or to
27 its parent the tree in the larger tree formed in the wide area network.

-14-

1 In addition, each computer forming a node in the multicast message delivery error recovery
2 tree 20, or in the larger tree formed in the wide area network, will receive at most only a limited
3 number of "NACK" messages, namely, at most a number corresponding to the number of its children
4 in the respective tree. As will be described below, in connection with establishment of the multicast
5 message delivery error recovery tree 20, each computer forming a node in the tree 20 can determine
6 the maximum number of children it can accommodate during establishment of the tree 20. In any
7 case, since none of the computers, including the source computer 11(s) for the multicast message
8 and the computers forming nodes in the multicast message delivery error recovery tree 20, or the
9 larger tree formed in the wide area network, receives an unduly large number of "NACK" messages
10 in the event of a failure to receive a multicast message, processing of the "NACK" messages will not
11 unduly negatively affect their other processing operations.

12 In one embodiment, the portion of the larger tree above the multicast message delivery error
13 recovery tree 20 established for the local area network 10 will be predetermined by, for example, a
14 system administrator, and one computer in the wide area network will be identified as a parent for
15 the computer 11(R) in the local area network 10 that is selected as the parent for the root node in the
16 tree 20. However, the computers 11(n) in the local area network 10 themselves determine the logical
17 organization of the multicast message delivery error recovery tree 20 for the local area network 10.
18 Operations performed by the computers 11(n) in establishing the multicast message delivery error
19 recovery tree 20 in the local area network 10 will be described below in connection with the flow
20 chart depicted in FIG. 4. Generally, the computers 11(n) make use of three types of messages,
21 namely, a node election message type, a node advertising message type, and a child solicitation
22 message type. The establishment of a multicast message delivery error recovery tree 20 proceeds
23 in a plurality of iterations, each iteration including two phases, namely, a node election phase and
24 a child solicitation phase. In the first iteration, a computer is selected as a root node in the tree 20,
25 and in each subsequent iteration, a number of additional computer(s) 11(n) will be added to the tree
26 20. In each iteration, the computers 11(n) use messages of the node election message type during
27 the node selection phase and messages of the node advertisement message type during the child

-15-

1 solicitation phase. The use of messages of the child solicitation message type will be described
2 below.

3 FIG. 3A depicts structure of a node election message 30 used by computers 11(n) in one
4 embodiment of the invention during the node election phase of each iteration. Generally, in each
5 iteration, each of the computers 11(n) that can become a node in the multicast message delivery error
6 recovery tree, referenced as "participating" in the iteration, multicast node election messages that
7 identify its suitability (as will be described below) of becoming a node in the tree. The node election
8 messages also identify a number of nodes that will be selected during the iteration. The computers
9 11(n) essentially self-select themselves as nodes at the end of the iteration. With reference to FIG.
10 3A, node election message 30 includes a header portion 31 and a data portion 32. The header
11 portion 31 contains message transfer protocol information, including a message type identifier,
12 source and destination addresses, and a time to live value. The message type identifier in the header
13 portion 31 identifies the message as being of the node election type, the source address identifies the
14 particular computer 11(n) that generated and transmitted the node election message and the
15 destination address identifies the multicast address for the computers 11(n) comprising the local area
16 network 10. The time to live value is selected to ensure that the node election messages are
17 transmitted only in the local area network.

18 The data portion 32 contains a number of fields that are used during the iteration in selecting
19 the computers 11(n) that are to be nodes in the tree 20, including a priority field 33, a capacity field
20 34, a computer system identifier field 35, a number of nodes field 36 and an election interval field
21 37. The priority field 33, capacity field 34 and computer system identifier field 35 contain values
22 that serve to rank the participating computers in their relative suitability to become a node during
23 the iteration. In particular, priority field contains a priority value that identifies a relative priority
24 for the computer 11(n) that generates the node election message to be a node. Illustrative priority
25 values used in priority field 33 include, for example, a highest "must be" a node value, a relatively
26 high "eager" to be a node value, and a relatively low "reluctantly would be" a node value. In
27 addition, the lowest "leaf only" node value can be assigned to a computer 11(n) if it will only operate

-16-

1 as a leaf in the multicast message delivery error recovery tree 20; in that case, the computer 11(n)
2 will not participate in the node election phase of any iteration. The capacity field 37 contains a child
3 capacity value that identifies the number of children the computer 11(n) can accommodate. The
4 computer system identifier field 35 contains a computer system identifier that identifies the computer
5 11(n) in the local area network 10; in one embodiment, the computer system identifier comprises the
6 network address for the computer 11(n).

7 The number of nodes field 36 contains a value that identifies the number of computers 11(n)
8 that are to be selected as nodes during the iteration. Thus, in the first iteration, in which the
9 computer to serve as the root node in the multicast message delivery error recovery tree 20 is
10 selected, the value contained in the number of nodes field 36 will be "one." If a subsequent iteration
11 is needed to select additional computers to operate as nodes in the multicast message delivery error
12 recovery tree 20, the computer 11(R) previously selected as the root node in the tree 20 will provide
13 a value identifying the number of computers 11(n) that are to be selected as nodes in the subsequent
14 iteration. These operations will continue through each subsequent iteration.

15 Finally, the election interval field 37 identifies the time interval between transmissions of
16 node election messages by the computer system which generates the respective message.

17 As noted above, the values in the priority field 33, capacity field 34 and computer system
18 identifier field 35 are used to rank the participating computers 11(n) as to their relative suitability
19 to be selected as nodes during the iteration. Generally, the ranking is based first on the priority
20 value, so that those computers 11(n) which have the higher priority values will be deemed more
21 suitable and those which have lower priority values will be deemed less suitable. As among
22 computers 11(n) which have the same priority value, suitability is determined based on their
23 respective capacity values, which identify the number of children that they can accommodate; that
24 is, computers 11(n) with higher capacity values will be deemed more suitable and computers 11(n)
25 with lower capacity values will be deemed less suitable, within the ranking as determined by their
26 priority values. Finally, as among computers 11(n) with the same priority and capacity values, the

-17-

1 suitability ranking is determined based on their respective computer system identifiers. Each
2 computer system identifier is unique, at least among computers 11(n) in the local area network, and
3 the use of the computer system identifier value is selected, somewhat arbitrarily, as a means to
4 provide rankings as among computers in the local area network 10 which have the same priority and
5 capacity values.

6 It will be appreciated that, if the priority values, capacity values and computer system
7 identifier values are all numerical values, with higher priority, capacity and computer system
8 identifier values each being represented by a higher numerical value and lower priority, capacity and
9 computer system identifier values each being represented by a lower numerical value, a computer
10 can generate a unitary suitability value "S" as <PRI_VAL|CAP_VAL|CSID_VAL>, where
11 "PRI_VAL" represents a field for the priority value, "CAP_VAL" represents a field for the capacity
12 value, "CSID_VAL" represents a field for the computer system identifier value, and "|" represents
13 the concatenation operation, with the respective fields used in the node election messages from all
14 of the computers 11(n) having the same number of bits. Each computer 11(n) determines its
15 suitability ranking in connection with the number of computers in the local area network 10 which
16 have higher suitability values than the computer 11(n).

17 FIG. 3B depicts the structure of a node advertising message 40 used in one embodiment of
18 the invention during the second phase of each iteration. In particular, the node advertising message
19 40 is used by computers 11(n) which have determined that they are nodes in the multicast message
20 delivery error recovery tree 20 to notify other computers in the local area network 10 that they are
21 available to have other computers logically connect to them as children in the tree 20. In that
22 connection, the computer 11(n) that transmits the node advertising message 40 also provides some
23 information to the other computers as to the status of the computer 11(n), which they (that is, the
24 other computers) can use to determine whether they should logically connect to that computer 11(n).

25 With reference to FIG. 3B, node advertising message 40 includes a header portion 41 and
26 a data portion 42. The header portion 41 contains message transfer protocol information, including

-18-

1 a message type identifier and source and destination addresses, and a time to live value. The
2 message type identifier in the header portion 41 identifies the message as being of the node
3 advertising type, the source address identifies the particular computer 11(n) that generated and
4 transmitted the node advertising message and the destination address identifies the multicast address.
5 The time to live value is selected to ensure that the node advertising messages are transmitted only
6 in the local area network.

7 The data portion 42 contains a number of fields that are used during the iteration in notifying
8 the other computers in the local area network 10 as to the logical connection availability of the
9 computer 11(n) that generates the message 40, including a priority field 43, a capacity field 44, a
10 computer system identifier field 45, an advertising interval field 46, a number of children field 47,
11 a connected field 48, a number of nodes field 49 and a status field 50. The priority field 43, capacity
12 field 44 and computer system identifier field 45 contain priority, capacity and computer system
13 identifier values that correspond to the respective values in fields 33, 34, and 35, respectively, of
14 node election message 30 as described above. The advertising interval field 46 identifies the time
15 interval between transmissions of a node advertising message 40 by the particular computer 11(n)
16 that generates the node advertising message 40.

17 The number of children field 47 of node advertising message 40 contains a value that
18 identifies the number of other computers to which the particular computer 11(n) is currently logically
19 connected when it (that is, the particular computer 11(n)) transmits the node advertising message 40.
20 A computer that receives a node advertising message 40 from the computer 11(n) can determine an
21 excess capacity value for computer 11(n), which is a measure of the ability of the computer 11(n)
22 to take on additional children, by determining the difference between the capacity value in field 44
23 and the number of children value in field 47. A computer may use the excess capacity value for
24 several purposes, including determining whether it should attempt to logically connect as a child to
25 a particular computer 11(n) based on its excess capacity in relation to the excess capacity values of
26 computers which form other nodes in the multicast message delivery error recovery tree 20. Other
27 uses of the excess capacity value will be described below.

-19-

1 The connected field 48 of node advertising message 40 contains a connected flag that, if set,
2 indicates that the computer 11(n) that generates the node advertising message 40 is logically
3 connected to multicast message delivery error recovery tree 20, and otherwise indicates that the
4 computer is not logically connected to the tree 20. The number of nodes field 49 identifies the
5 current number of nodes that have been selected for the tree 20. The status field 50 provides
6 information as to the computer's status in the tree 20. In one embodiment the status field 50
7 normally has an "active" value, which indicates that the computer 11(n) that generates the node
8 advertising message 40 is currently a node in the tree 20. However, as will be described below, the
9 status field 50 can also have a "resigning" value, which indicates that the computer 11(n) that
10 generates the node advertising message 40 is a node in tree 20, but is in the process of resigning as
11 a node and converting to a leaf. If status field 50 of the node advertising message 40 transmitted
12 by a computer 11(n) indicates the "resigning" status, its children in the tree will need to attempt to
13 logically connect to other computers that are nodes in the tree.

14 As noted above, establishment of multicast message delivery error recovery tree 20 proceeds
15 in a plurality of iterations. Generally, the first iteration will start when at least one of the computers
16 11(n) of the local area network 10 has been powered on and initialized. After a computer 11(n) has
17 been initialized, if it is associated with a priority value that would allow it to be a node in the
18 multicast message delivery error recovery tree 20, that is, if it is associated with a priority value other
19 than "leaf only," it will begin periodically transmitting node election messages 30. If, within a
20 selected time interval after beginning transmitting node election messages 30, the computer 11(n)
21 does not receive a node election message 30 from other computers in the network, it (that is,
22 computer 11(n)) will determine that it is "elected" as a node in the multicast message delivery error
23 recovery tree 20, and stop transmitting node election messages 30. Since the computer 11(n)
24 comprising the elected node is, at this point, the only node in the tree 20, it (that is, the computer
25 11(n)) will comprise the root node 11(R) in the tree 20. After the computer 11(n) determines that
26 it is the elected node, will begin transmitting node advertising messages 40 to notify other computers
27 in the local area network 10 of its availability to accept children. If other computers in the local area

-20-

1 network 10 that have been powered on and initialized, after receiving a node advertising message
2 40, they can communicate with the computer 11(n) to attempt to logically become children of the
3 computer 11(n) in the tree 20. It will be appreciated that this may occur (that is, that the computer
4 11(n) does not receive any node election messages 30 from other computers in the local area network
5 within the selected time interval) if there are no other computers in the network which have been
6 powered on and initialized which can operate as a node in the multicast message delivery error
7 recovery tree 20 for the local area network. In addition, after being "elected" a node in the tree 20,
8 the computer 11(n) can increment its priority value by a predetermined "elected node" increment
9 value, the purpose for which will be described below.

10 On the other hand, if, within the selected time interval, the computer 11(n) receives a node
11 election message 30 from at least one other computer 11(n') in the local area network 10, it will
12 compare the suitability value (that is, as described above, the concatenation of the priority value,
13 capacity value and computer system identifier value from fields 33, 34 and 35) from the received
14 node election message to its own suitability value (that is, the concatenation of its priority value,
15 capacity value and computer system identifier value) that it is providing in the node election
16 messages 30 that it is transmitting. If the computer 11(n) determines that its suitability value is less
17 than the suitability value in the received node election message, it will stop transmitting node
18 election messages. On the other hand, if the computer 11(n) determines that its suitability value is
19 greater than the suitability value in the received node election message, it will continue transmitting
20 node election messages. The other computer 11(n') will perform similar operations, so that, if it
21 determines that its suitability value is higher than that of computer 11(n), it will continue
22 transmitting node election messages, but if it determines that its suitability value is lower than that
23 of computer 11(n), it will stop transmitting node election messages.

24 Each computer in the local area network with a priority level above "leaf only" will perform
25 similar operations. At the end of the selected time interval, one of the computers, if it is still
26 transmitting node election messages, then it has not received a node election message from a
27 computer in the local area network which has a higher suitability value, and so it (that is, the still-

-21-

1 transmitting computer) will determine that it is "elected" a node, in this case the root node, in the tree
2 20, and therefore will constitute the computer 11(R) (reference FIG. 2) in the tree 20. After
3 computer 11(R) determines that it is the elected node, it will stop transmitting node election
4 messages 30 and begin transmitting node advertising messages 40 to notify other computers in the
5 local area network 10 of its availability to accept children. Other computers in the local area network
6 10 that have been powered on and initialized can, after receiving a node advertising message 40,
7 communicate with the computer 11(R) to attempt to become children of the elected computer in the
8 tree 20. In addition, the computer 11(R) can increment its priority value by the predetermined
9 "elected node" increment value.

10 If the computer 11(R) that forms the root node of the multicast message delivery error
11 recovery tree 20 determines, during the communications with the other computers in the local area
12 network 10, that it has the capacity to accommodate all of the other computers as leaves, it can
13 logically connect to the other computers at that point. On the other hand, if the computer 11(R)
14 determines that more computers are requesting to become children thereof than it can accommodate,
15 it (that is, the computer 11(R)) can initiate a second iteration in the tree establishment process, to
16 enable election of additional nodes for the multicast message delivery error recovery tree 20. It will
17 be appreciated that the additional nodes will comprise one or more levels in the tree 20 below the
18 root level. To initiate the second iteration, the computer 11(R) will resume transmitting node
19 election messages 30. In this iteration, the node election messages 30 generated by the computer
20 11(R) will contain a number of nodes value in field 36 that is greater than one by an amount
21 corresponding to the number of new nodes that are to be elected. In the node election messages
22 generated by the computer 11(R), the priority value in field 33 will reflect the incremented priority
23 value, that is, its original priority value incremented by the predetermined "elected node" increment
24 value.

25 Each of the other computers in the local area network 10 which can also become nodes in
26 the tree (that is, each computer other than computer 11(R) whose priority value is greater than "leaf
27 only") will also resume transmitting node election messages 30. In this case, each of the other

-22-

1 computers will use their own priority, capacity and computer system identifier values in fields 33
2 through 35 of their node election messages, but will use the number of nodes and election interval
3 values from fields 36 and 37 in the node election messages 30 received from the computer 11(R) in
4 node election messages that they transmit. While transmitting node election messages, each of the
5 computers will also receive node election messages from the other computers and will compare the
6 suitability values therefrom with their own suitability values. If a computer that is transmitting node
7 election messages 30, determines that it has received node election messages 30 with higher
8 suitability values than its (that is, the computer's) suitability value, from a number of other computers
9 corresponding to the number of nodes value in field 36, it (that is, the computer) will stop sending
10 node election messages 30. Until that occurs, the computer will continue transmitting node election
11 messages until the end of the selected time interval. The number of computers which are continuing
12 to transmit node selection messages 30 at the end of the selected time interval will correspond to the
13 number of nodes specified in the number of nodes field 36, and so the computers that are
14 transmitting node election messages 30 at the end of the selected time interval will constitute the set
15 of computers that are elected to be nodes in the tree-20, including the root node.

16 It will be appreciated that the computer 11(R) that was elected the root node during the
17 previous iteration may, but need not, have the highest suitability value (even with the priority value
18 incremented by the predetermined "elected node" increment value) as among the elected nodes since
19 one or more computers associated with higher suitability values than the computer 11(R) may have
20 been powered on and initialized and begun transmitting node election messages. If the computer
21 11(R) that was elected as the root node during the previous iteration has a priority level, as
22 incremented by the "elected node" increment value, that is higher than the other computers, then it
23 will remain the root node. However, if another computer has a priority level that is sufficiently high
24 that its suitability value is higher than the suitability of computer 11(R), the other computer will be
25 elected root node during the current iteration. Indeed, if sufficient numbers of newly-transmitting
26 computers have higher suitability values than the previously elected computer 11(R), that computer
27 11(R) may be "de-elected," that is, it may not be among the elected nodes for the iteration. In that

-23-

1 case, the computer with the highest suitability value will be elected as the new root node 11(R) for
2 subsequent operations, and the other elected computers will form other portions, including (but not
3 necessarily limited to) the second level in the tree 20. However, it will be appreciated that, by
4 incrementing the priority level of computer 11(R) that was previously-elected the root node in the
5 tree 20, by the "elected node" increment value, the previous election will be "sticky," that is, it will
6 not be disturbed unless the priority level of the other computer(s) is or are sufficiently higher than
7 the original, non-incremented, priority level of the computer 11(R) to warrant disturbing the previous
8 election.

9 Thereafter, the elected computers will transmit node advertising messages 40, as in the first
10 iteration, to notify other computers in the local area network 10 of their availability to accept
11 children. The computers elected during the iteration, other than the computer 11(R) that forms the
12 root node for the tree 20, will receive the node advertising messages 40 transmitted by the computer
13 11(R) and will communicate with it to become logically connected thereto, thereby to establish them
14 as forming the second level of the tree. If a newly-elected node cannot logically connect to the
15 computer 11(R), it may logically connect to another newly-elected computer to begin the third level
16 in the tree. The other, non-elected computers will communicate with any of the elected computers
17 to negotiate formation of logical connections therewith.

18 If the root node determines that additional nodes of the tree are required, it can repeat the
19 operations through subsequent iterations. In each iteration, the computers newly-elected to form
20 node(s) in the tree 20 will increment their priority values by the predetermined "elected node" value,
21 and, if a previously-elected computer is "de-elected," it will reduce its priority value to the original
22 priority value. These operations will continue until preferably all of the computers 11(n) in the local
23 area network 10 form part of the multicast message delivery error recovery tree 20, either as nodes
24 or leaves. In addition, the root node 11(R) will logically connect as a child to a node specified for
25 the local area network in the tree for the wide area network. After the multicast message delivery
26 error recovery tree 20 has been established, the computers 11(n) comprising nodes in the tree will

1 continually transmit node advertising messages 40, so that newly-initialized computers can identify
2 nodes in the multicast message delivery error recovery tree 20 to which they may logically connect.

3 As noted above, a computer 11(n) in the local area network, after it is powered on and
4 initialized, if it has a priority value that is other than "leaf only," will begin transmitting node
5 election messages. If a multicast message delivery error recovery tree 20, or any portion thereof
6 established as described above, already exists, the computers 11(n) will repeat the tree-establishment
7 operations as described above. In those operations, the computers 11(n) which form nodes in the
8 tree will use their priority values, as incremented by the "elected node" increment value, so that the
9 tree 20 will not be disturbed during the new tree-establishment operations unless the newly-
10 initialized computer has a sufficiently high priority value. If the newly-initialized computer has a
11 priority value that is higher than the priority value associated with the computer 11(R) forming the
12 root node for the tree 20, it (that is, the newly-initialized computer) will become the new root node
13 for the tree. In that case, the computer 11(R') forming the new root node will logically connect as
14 a child to a node specified for the local area network in the tree for the wide area network, and the
15 computer 11(R) will logically connect to the computer 11(R') forming the new root node.

16 In addition, after the multicast message delivery error recovery tree 20 has been established
17 in the local area network, the computers 11(n) will periodically repeat the election process. In that
18 operation, any of the computers 11(n), other than a computer whose priority level is "leaf only," can
19 transmit a node election message. The node election message transmitted by a computer 11(n) will
20 include computer's suitability value, including its priority level either as incremented by the "elected
21 node" increment value if the computer currently forms a node in the tree 20, or not incremented if
22 the computer does not form a node in the tree 20. When other computers receive the node election
23 message, if they have not themselves already initiated resuming transmission of respective node
24 election messages, will resume transmission, and the node election process as described above will
25 be repeated.

-25-

1 In one embodiment, after the multicast message delivery error recovery tree 20 has been
2 established, the computers 11(n) that form nodes in the tree 20 attempt to optimize the tree. In that
3 case, the computer 11(n_{LSV}) that forms the node in the tree 20 with the lowest suitability value will
4 determine whether it should continue operating as a node, or enable any children logically connected
5 thereto to logically connect to computers that form other nodes in the tree. To accomplish that, the
6 computer 11(n_{LSV}) will determine whether excess capacity of all of the other computers forming
7 nodes in the tree 20, is greater than or equal to the number of its children. In that case, the computer
8 11(n_{LSV}) will transmit node advertising messages 40 that identify its status as "resigning." When the
9 children of computer 11(n_{LSV}) receive such node advertising messages, they will attempt to logically
10 connect to computers that form the other nodes in the tree 20. After each such child logically
11 connects to another computer, it will sever the logical connection with the computer 11(n_{LSV}), and
12 the computer 11(n_{LSV}) can reduce the number of children value in the node advertising messages
13 transmitted thereby. Between the time at which the child has logically connected to the other
14 computer and the time at which it (that is, the child) has severed the logical connection with the
15 computer 11(n_{LSV}), it (that is, the child) will effectively be connected to both computers, and may
16 transmit "NACK" messages to either or both computers. This will ensure that all computers 11(n)
17 are logically connected in the multicast message delivery error recovery tree 20. After all of the
18 children of computer 11(n_{LSV}) have severed their connections with computer 11(n_{LSV}), that computer
19 11(n_{LSV}) will comprise a leaf in the multicast message delivery error recovery tree 20 and will stop
20 transmitting node advertising messages 40. These operations can be repeated through a plurality of
21 iterations, for successive computers that are nodes with low suitability values.

22 With this background, the operations performed by a computer 11(n) in connection with the
23 invention will be described in connection with FIGS. 4A through 4G. Generally, operations
24 described in connection with FIG. 4 include several sequences, including

25 (i) a node election message transmit sequence (FIGS. 4A and 4B) depicting operations
26 performed by a node election message transmitter maintained by the computer 11(n) in connection
27 with transmission of node election messages;

-26-

1 (ii) a node election message receive sequence (FIGS. 4C and 4D), depicting operations
2 performed by a node election message receiver maintained by the computer 11(n) in connection with
3 reception of node election messages,

4 (iii) a tree organization sequence (FIG. 4E through 4G), depicting operations performed by
5 a tree organization component maintained by the computer 11(n) in connection with logically
6 connecting the computer 11(n) to a respective parent and children as appropriate

7 which the computer 11(n) performs during an iteration, as described above.

8 As described above, a computer 11(n) will begin transmitting node election messages 30
9 either

10 (i) after being powered up and initialized,

11 (ii) a selected time period after a multicast message delivery error recovery tree 20 was
12 previously established, or

13 (iii) in response to receipt of a node election message from another computer,

14 for a particular node election message time period. Thus, with reference to FIG. 4A, the computer
15 11(n) will, after being powered up and initialized (step 100), initially determine whether it has a
16 priority level greater than "leaf only" (step 101). If the computer 11(n) makes a negative
17 determination in step 101, that is, if it has a priority level of "leaf only," it will operate only as a leaf
18 in a multicast message delivery error recovery tree 20 established for local area network 10. In that
19 case, the computer 11(n) can wait for a node advertising message 40 from a computer that forms a
20 node in a multicast message delivery error recovery tree 20, or it can begin transmitting child
21 solicitation messages to enable computers that form nodes in a tree 20 to transmit node advertising
22 messages 40.

23 On the other hand, if the computer 11(n) makes a positive determination in step 101, that is,
24 if it has a priority level above "leaf only," it can operate as a node in a multicast message delivery

-27-

1 error recovery tree 20 established in local area network 10. In that case, computer 11(n) will proceed
2 to a series of steps to initiate transmission of node election messages 30. In that operation, the
3 computer 11(n) will establish and initialize a node election message period timer (step 102) to be
4 used to time the time period over which the computer 11(n) will transmit node election messages.
5 In addition, the computer 11(n) will establish and initialize a number of nodes store (step 103),
6 which stores a number of nodes value that will be used in the node election messages, and a node
7 counter (step 104) that will be used to count the number of computers 11(n) from which it receives
8 node election messages 30 with higher suitability values than its own suitability value. Initially the
9 node counter will be set to a value of "zero." Following step 104, the computer 11(n) will determine
10 whether the number of nodes value stored in the number of nodes store is higher than the value
11 provided by the node counter (step 105).

12 If the computer 11(n) makes a positive determination in step 105, it will transmit a node
13 election message including its priority, capacity and computer system identifier values in fields 33,
14 34 and 35, the number of nodes value from the number of nodes store in field 36, and the election
15 interval value in field 37 (step 106). Following step 106, computer 11(n) will establish and initialize
16 a node election message transmission interval timer to time the interval between transmission of
17 node election messages, which will initially be set to a default value established for the computer
18 11(n) (step 107). When the computer 11(n) determines that the node election message transmission
19 interval timer established in step 107 has timed out (step 108), if it determines that the node election
20 message period timer has not timed out (step 109), the computer 11(n) will return to step 104 to
21 determine whether to transmit another node election message 30. It will be appreciated that the
22 computer 11(n) will repeat steps 104 through 109 until it determines

23 (i) in step 105 that the value provided by node counter is greater than or equal to the number
24 of nodes value, or

25 (ii) in step 108 that the node election message period timer has timed out.

-28-

1 If the computer 11(n) determines in step 105 that the value provided by node counter is greater than
2 or equal to the number of nodes value, then it has received node election messages from a number
3 of computers in the local area network whose suitability values are greater than the suitability value
4 of computer 11(n) corresponding to the number of nodes to be established in the multicast message
5 delivery error recovery tree 20, and so it (that is, computer 11(n)) can stop transmitting node election
6 messages 30.

7 If the computer 11(n) receives a node election message 30 from another computer 11(n') in
8 the local area network 10 (step 120, FIG. 4C), it will initially determine whether it (that is, the
9 computer 11(n)) is currently also transmitting node election messages (step 121). If the computer
10 11(n) makes a negative determination in step 121, that is, if it determines that it is not currently
11 transmitting node election messages, if its priority level is greater than "leaf only," the computer will
12 start transmitting node election messages as described above in connection with FIGS. 4A and 4B
13 (step 122). In addition, the computer 11(n) will determine whether the number of nodes value in
14 field 36 of the received message is larger than the number of nodes value in its number of nodes
15 store (step 123). If the computer 11(n) makes a positive determination in step 123, that is, if it
16 determines that the number of nodes value in field 36 of the received message is larger than the
17 number of nodes value in its number of nodes store, then the computer 11(n') from which the node
18 election message 30 was received forms a part of a previously-established multicast message
19 delivery error recovery tree 20 that include a plurality of nodes. In that case, the computer 11(n) will
20 increase the number of nodes value in its number of nodes store (step 124) and adjust its node
21 election message period time interval to be used when the node election message transmission
22 interval timer is next established in step 107 (step 125) for use during transmission of node election
23 messages 30. In addition, the computer 11(n) will determine whether the suitability value of the
24 computer 11(n') from which the message was received is greater than its suitability value (step 126).
25 If the computer 11(n) makes a positive determination in step 126, then it determines whether it
26 previously received a node election message 30 from computer 11(n') during the node election

-29-

1 message transmission interval (step 127) and, if so, will increment the node counter used to control
2 node election message transmission (step 128).

3 Returning to step 109 (FIG. 4B), after the node election message period timer has timed out,
4 the computer 11(n) will perform operations, depicted in FIG. 4E, in connection with logically
5 connecting to other computers as parent and/or children as appropriate. Thus, after the node election
6 message period timer has timed out, the computer 11(n) will first determine whether it is the root
7 node in the tree (step 140). In that operation, the computer 11(n) can determine whether it is the root
8 node by determining whether the node counter has the value zero, which would indicate that it (that
9 is, the computer 11(n) did not receive node election messages 30 from any other computer in the
10 local area network with higher suitability values than that of computer 11(n). If the computer 11(n)
11 makes a positive determination in step 140, it will attempt to form a logical connection to the parent
12 assigned to the local area network in the portion of the tree formed therefor in the wide area network
13 (step 141).

14 On the other hand, if the computer 11(n) makes a negative determination in step 140, it will
15 determine whether it is a node other than the root node in the multicast message delivery error
16 recovery tree 20. In that operation, the computer 11(n) can determine whether the number of nodes
17 value stored in the number of nodes store is higher than the value provided by the node counter (step
18 142). If the computer 11(n) makes a positive determination in step 142, then it will form a node in
19 the tree; on the other hand, if the computer 11(n) makes a negative determination in step 140, then
20 it received node election messages 30 from at least a number of other computers in the local area
21 network 10 corresponding to the number of nodes value identifying the number of nodes to be
22 elected.

23 Following step 141, or step 142 if a positive determination is made in that step, the computer
24 11(n) will first increase its priority value by the "elected node" value (step 143). Thereafter, the
25 computer 11(n) will begin transmitting node advertising messages to identify the availability of the
26 computer 11(n) to logically connect to other computers as children (step 144). In addition, if the

-30-

1 computer 11(n) determined in step 140 that it is not the root node, it will respond to the node
2 advertising messages generated by other computers which have priority, capacity and computer
3 system identifier values in fields 43-45 which define suitability values greater than its own to attempt
4 to logically connect to them as children, thereby to establish respective levels in the tree 20
5 (step 145). After the computer 11(n) has logically connected to another computer as parent of the
6 other computer, it will increase the number of children value in field 47 of node advertising
7 messages transmitted thereby (step 146). On the other hand, after the computer 11(n) has been
8 accepted by another computer as a child of the other computer, it will be logically connected to that
9 other computer and can save the identification of the other computer for use in sending "NACK"
10 messages, and set the connected flag 48 in node advertising messages 40 transmitted thereby (step
11 147).

12 After the computer 11(n) determines that the multicast message delivery error recovery tree
13 20 has been established (step 148), it will determine whether its suitability value is the lowest among
14 the computers elected nodes in the tree (step 149). The computer 11(n) can do that by, for example,
15 determining whether the value provided by its node counter corresponds to the number of nodes
16 value in field 49 of the node advertising messages 40 transmitted by the computers comprising nodes
17 in the tree 20. If the computer 11(n) makes a positive determination in step 149, it will determine
18 whether the excess capacity, as described above, of the computers comprising the other nodes is
19 greater than the number of computers logically connected to it as children (step 150). If the
20 computer 11(n) makes a positive determination in step 150, if it has any computers logically
21 connected thereto as children (step 151) it will provide a value of "resigning" in the status field 50
22 of node advertising messages 40 transmitted thereby to enable its children to attempt to logically
23 connect to other computers forming nodes in the tree 20 (step 152). As the computer 11(n) receives
24 notification from each child that logically connects to another computer forming a node in the tree
25 20 (step 153), it will reduce the number of children indicated in field 47 of node advertising
26 messages transmitted thereby (step 154), and, after it determines that the number of children has

-31-

1 reduced to zero (step 155), it will stop transmitting node advertising messages (step 156) and
2 decrease its priority value by the "elected node" value (step 157).

3 It will be appreciated that a system in accordance with the invention can be constructed in
4 whole or in part from special purpose hardware or a general purpose computer system, or any
5 combination thereof, any portion of which may be controlled by a suitable program. Any program
6 may in whole or in part comprise part of or be stored on the system in a conventional manner, or it
7 may in whole or in part be provided in to the system over a network or other mechanism for
8 transferring information in a conventional manner. In addition, it will be appreciated that the system
9 may be operated and/or otherwise controlled by means of information provided by an operator using
10 operator input elements (not shown) which may be connected directly to the system or which may
11 transfer the information to the system over a network or other mechanism for transferring
12 information in a conventional manner.

13 The foregoing description has been limited to a specific embodiment of this invention. It will
14 be apparent, however, that various variations and modifications may be made to the invention, with
15 the attainment of some or all of the advantages of the invention. It is the object of the appended
16 claims to cover these and such other variations and modifications as come within the true spirit and
17 scope of the invention.

18 What is claimed as new and desired to be secured by Letters Patent of the United States is:

-32-

CLAIMS

1 1. A method of logically organizing, in a digital data network comprising a plurality of devices
2 interconnected by a communication link, the devices into a tree structure, each of the devices having
3 an associated suitability value, the method comprising:

4 A. a node election step in which at least one of the devices is enabled to broadcast, over the
5 communication link, at least one node election message including the respective device's
6 suitability value, any devices broadcasting node election messages also being enabled to
7 receive from the communication link any node election messages broadcast by other devices
8 and to determine whether it is elected a node in the tree structure in connection with a
9 comparison between its suitability value and suitability values in any received node election
10 messages; and

11 B. a tree establishment step in which at least one device which is elected a node in the tree
12 structure is enabled to communicate with at least one other device to facilitate the other
13 device becoming a child of the device that is elected a node in the tree.

1 2. A method as defined in claim 1 in which the node election step and the tree establishment step are
2 performed in at least one iteration, in said at least one iteration, during the node election step one
3 device being elected a node, as a root node, in the tree structure.

1 3. A method as defined in claim 2 in which the node election step and the tree establishment step are
2 performed in a series of iterations, in a first iteration the root node being elected during the node
3 election step, and each subsequent iteration at least one other device being elected as a node in the
4 tree structure.

-33-

1 4. A method as defined in claim 3 in which the at least one other device elected as a node in the tree
2 structure, during the tree establishment step of at least one of said iterations, being enabled to
3 communicate with the device elected as root node to facilitate becoming a child thereof.

1 5. A method as defined in claim 3 in which the device elected as the root node is enabled to
2 determine the number of additional nodes to be elected during each subsequent iteration.

1 6. A method as defined in claim 5 in which the device elected as the root node transmits, in its at
2 least one node election message transmitted during each iteration after the first iteration, a number
3 of nodes value identifying the number of additional nodes to be elected during the respective
4 iteration.

1 7. A method as defined in claim 6 in which each of the devices enabled to broadcast at least one node
2 election message transmits, during each iteration after the first iteration, a number of nodes value
3 corresponding to the number of nodes value transmitted by the device elected as the root node.

1 8. A method as defined in claim 1 in which each device is assigned a priority value, the respective
2 device's suitability value being a function of its priority value.

-34-

1 9. A method as defined in claim 1 in which each device has an associated capacity value associated
2 with a number of children that it can accommodate, the respective device's suitability value being
3 a function of its priority value.

1 10. A method as defined in claim 1 in which said at least one device is enabled to begin broadcasting
2 at least one node election message after it is initialized.

1 11. A method as defined in claim 1 in which said at least one device is enabled to broadcast a
2 plurality of node election messages, the at least one device stopping broadcasting node election
3 messages if it receives a node election message from another device having a higher suitability value
4 than the suitability value associated with the at least one device.

1 12. A method as defined in claim 1 in which, during the tree establishment step the at least one
2 device that is elected a node in the tree is enabled to broadcast a node advertising message over the
3 communication link indicating that it is available as a node in the tree.

1 13. A method as defined in claim 12 in which the at least one device that is elected a node in the tree
2 is enabled to broadcast a node advertising message in response to receipt of a child solicitation
3 message transmitted by at least one other device.

1 14. A method as defined in claim 12 in which at least one other device, after receiving a node
2 advertising message, communicates with the at least one device that is elected a node in the tree to
3 facilitate becoming a child thereof.

-35-

1 15. A system for logically organizing devices, in a digital data network comprising a plurality of
2 devices interconnected by a communication link, into a tree structure, each of the devices having an
3 associated suitability value, the system comprising:

4 A. a node election element configured to enable at least one of the devices to broadcast, over
5 the communication link, at least one node election message including the respective device's
6 suitability value, and to enable any devices broadcasting node election messages to receive
7 from the communication link any node election messages broadcast by other devices and to
8 determine whether it is elected a node in the tree structure in connection with a comparison
9 between its suitability value and suitability values in any received node election messages;
10 and

11 B. a tree establishment element configured to enable at least one device which is not elected a
12 node in the tree structure to communicate with at least one device which is elected a node
13 in the tree structure to facilitate becoming a child of the device that is elected a node in the
14 tree.

1 16. A system as defined in claim 15 in which the node election element and the tree establishment
2 element are configured to operate in at least one iteration, in said at least one iteration, during the
3 node election step one device being elected a node, as a root node, in the tree structure.

1 17. A system as defined in claim 16 in which the node election element and the tree establishment
2 element are configured to operate in a series of iterations, in a first iteration the root node being
3 elected during the node election step, and each subsequent iteration at least one other device being
4 elected as a node in the tree structure.

-36-

1 18. A system as defined in claim 17 in which the node election element is configured to enable at
2 least one other device elected as a node in the tree structure, during the tree establishment step of at
3 least one of said iterations, to communicate with the device elected as root node to facilitate
4 becoming a child thereof.

1 19. A system as defined in claim 17 in which the node election element enables the device elected
2 as the root node to determine the number of additional nodes to be elected during each subsequent
3 iteration.

1 20. A system as defined in claim 19 in which the node election element enables the device elected
2 as the root node to transmit, in its at least one node election message transmitted during each
3 iteration after the first iteration, a number of nodes value identifying the number of additional nodes
4 to be elected during the respective iteration.

1 21. A system as defined in claim 20 in which the node election element enables each of the devices
2 to broadcast at least one node election message transmits, during each iteration after the first
3 iteration, a number of nodes value corresponding to the number of nodes value transmitted by the
4 device elected as the root node.

1 22. A system as defined in claim 15 in which each device is assigned a priority value, the respective
2 device's suitability value being a function of its priority value.

-37-

1 23. A system as defined in claim 15 in which each device has an associated capacity value associated
2 with a number of children that it can accommodate, the respective device's suitability value being
3 a function of its priority value.

1 24. A system as defined in claim 15 in which said node election element enables at least one device
2 to begin broadcasting at least one node election message after it is initialized.

1 25. A system as defined in claim 15 in which said node election element enables at least one device
2 to broadcast a plurality of node election messages, and to stop broadcasting node election messages
3 if it receives a node election message from another device having a higher suitability value than the
4 suitability value associated with the at least one device.

1 26. A system as defined in claim 15 in which the tree establishment element enables the at least one
2 device that is elected a node in the tree to broadcast a node advertising message over the
3 communication link indicating that it is available as a node in the tree.

1 27. A system as defined in claim 26 in which the tree establishment element enables at least one
2 device that is elected a node in the tree to broadcast a node advertising message in response to receipt
3 of a child solicitation message transmitted by at least one other device.

1 28. A system as defined in claim 26 in which the tree establishment element enables at least one
2 other device, after receiving a node advertising message, to communicate with the at least one device
3 that is elected a node in the tree to facilitate becoming a child thereof.

-38-

1 29. A computer program product for use in connection with a computer, the computer program
2 product enabling said computer to, in a digital data network comprising a plurality of computers
3 interconnected by a communication link, organize themselves into a tree structure, each of the
4 computers having an associated suitability value, the computer program product comprising a
5 computer-readable medium having encoded thereon:

6 A. a node election module configured to enable at least one of the computers to broadcast, over
7 the communication link, at least one node election message including the respective
8 computer's suitability value, and to enable any computers broadcasting node election
9 messages to receive from the communication link any node election messages broadcast by
10 other computers and to determine whether it is elected a node in the tree structure in
11 connection with a comparison between its suitability value and suitability values in any
12 received node election messages; and

13 B. a tree establishment module configured to enable at least one computer which is not elected
14 a node in the tree structure to communicate with at least one computer which is elected a
15 node in the tree structure to facilitate becoming a child of the computer that is elected a node
16 in the tree.

1 30. A computer program product as defined in claim 29 in which the node election module and the
2 tree establishment module are configured to operate in at least one iteration, in said at least one
3 iteration, during the node election step one computer being elected a node, as a root node, in the tree
4 structure.

-39-

1 31. A computer program product as defined in claim 30 in which the node election module and the
2 tree establishment module are configured to operate in a series of iterations, in a first iteration the
3 root node being elected during the node election step, and each subsequent iteration at least one other
4 computer being elected as a node in the tree structure.

1 32. A computer program product as defined in claim 31 in which the node election module is
2 configured to enable at least one other computer elected as a node in the tree structure, during the
3 tree establishment step of at least one of said iterations, to communicate with the computer elected
4 as root node to facilitate becoming a child thereof.

1 33. A computer program product as defined in claim 31 in which the node election module enables
2 the computer elected as the root node to determine the number of additional nodes to be elected
3 during each subsequent iteration.

1 34. A computer program product as defined in claim 33 in which the node election module enables
2 the computer elected as the root node to transmit, in its at least one node election message
3 transmitted during each iteration after the first iteration, a number of nodes value identifying the
4 number of additional nodes to be elected during the respective iteration.

1 35. A computer program product as defined in claim 34 in which the node election module enables
2 each of the computers to broadcast at least one node election message transmits, during each
3 iteration after the first iteration, a number of nodes value corresponding to the number of nodes value
4 transmitted by the computer elected as the root node.

-40-

1 36. A computer program product as defined in claim 29 in which each computer is assigned a
2 priority value, the respective computer's suitability value being a function of its priority value.

1 37. A computer program product as defined in claim 29 in which each computer has an associated
2 capacity value associated with a number of children that it can accommodate, the respective
3 computer's suitability value being a function of its priority value.

1 38. A computer program product as defined in claim 29 in which said node election module enables
2 at least one computer to begin broadcasting at least one node election message after it is initialized.

1 39. A computer program product as defined in claim 29 in which said node election module enables
2 at least one computer to broadcast a plurality of node election messages, and to stop broadcasting
3 node election messages if it receives a node election message from another computer having a higher
4 suitability value than the suitability value associated with the at least one computer.

1 40. A computer program product as defined in claim 20 in which the tree establishment module
2 enables the at least one computer that is elected a node in the tree to broadcast a node advertising
3 message over the communication link indicating that it is available as a node in the tree.

1 41. A computer program product as defined in claim 40 in which the tree establishment module
2 enables at least one computer that is elected a node in the tree to broadcast a node advertising

-41-

3 message in response to receipt of a child solicitation message transmitted by at least one other
4 computer.

1 42. A computer program product as defined in claim 40 in which the tree establishment module
2 enables at least one other computer, after receiving a node advertising message, to communicate
3 with the at least one computer that is elected a node in the tree to facilitate becoming a child thereof.

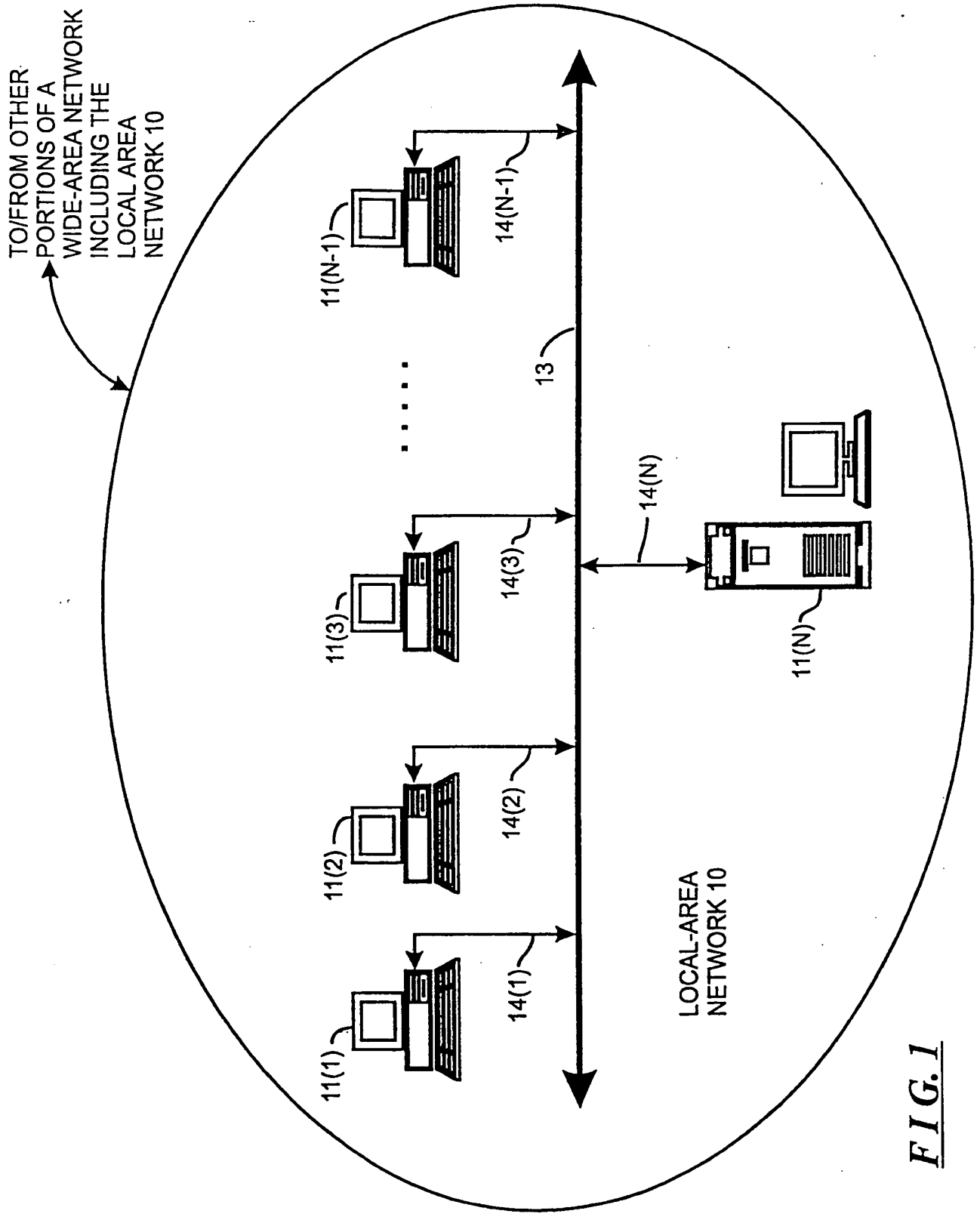
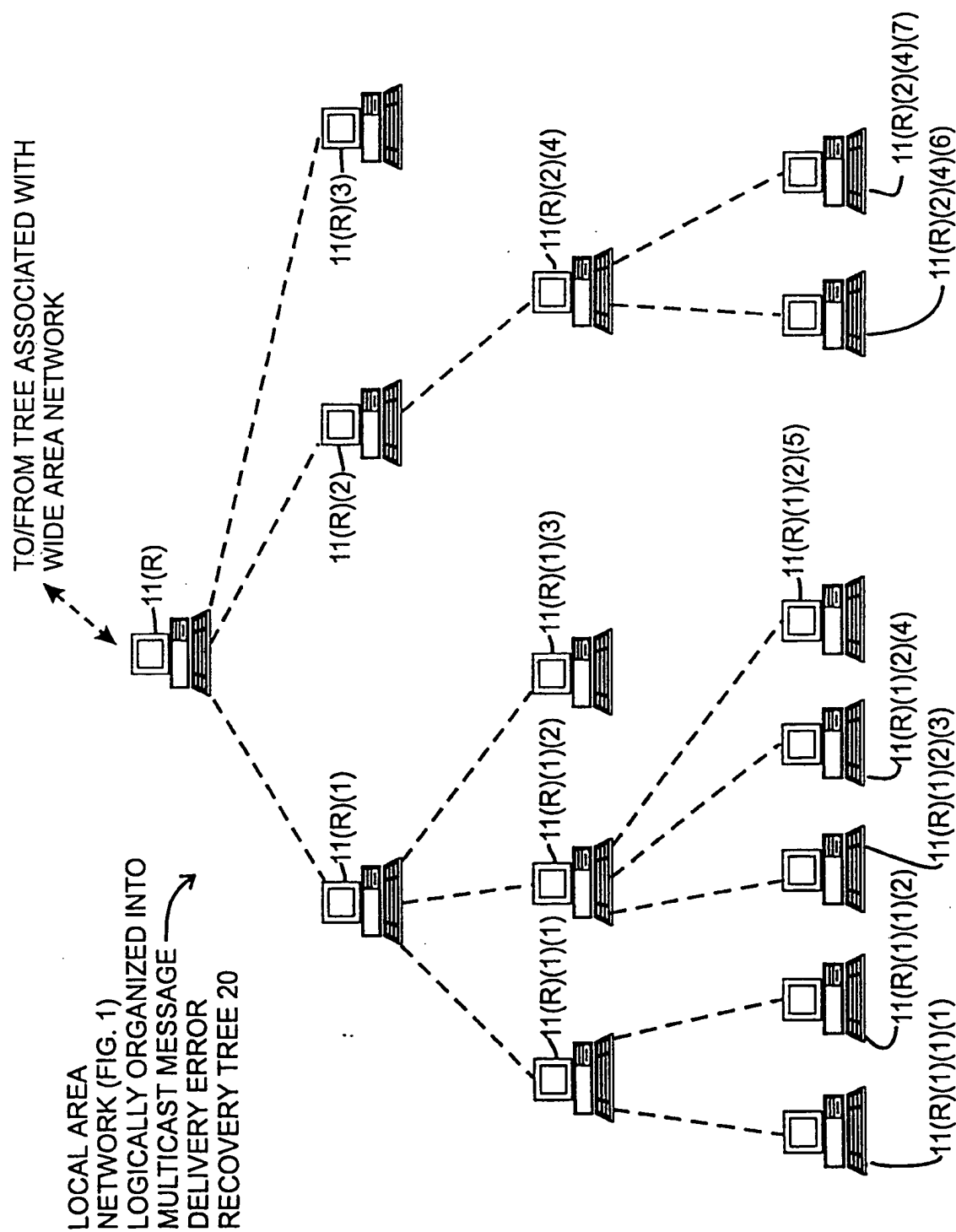
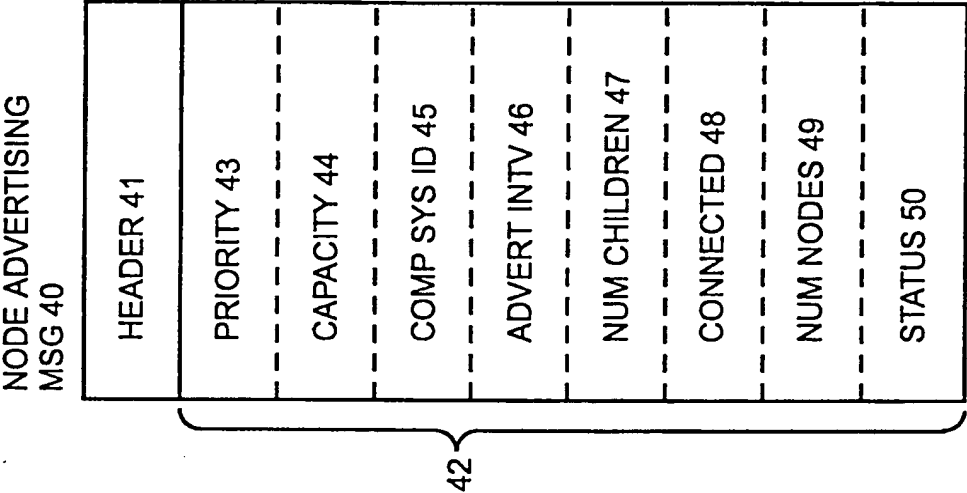
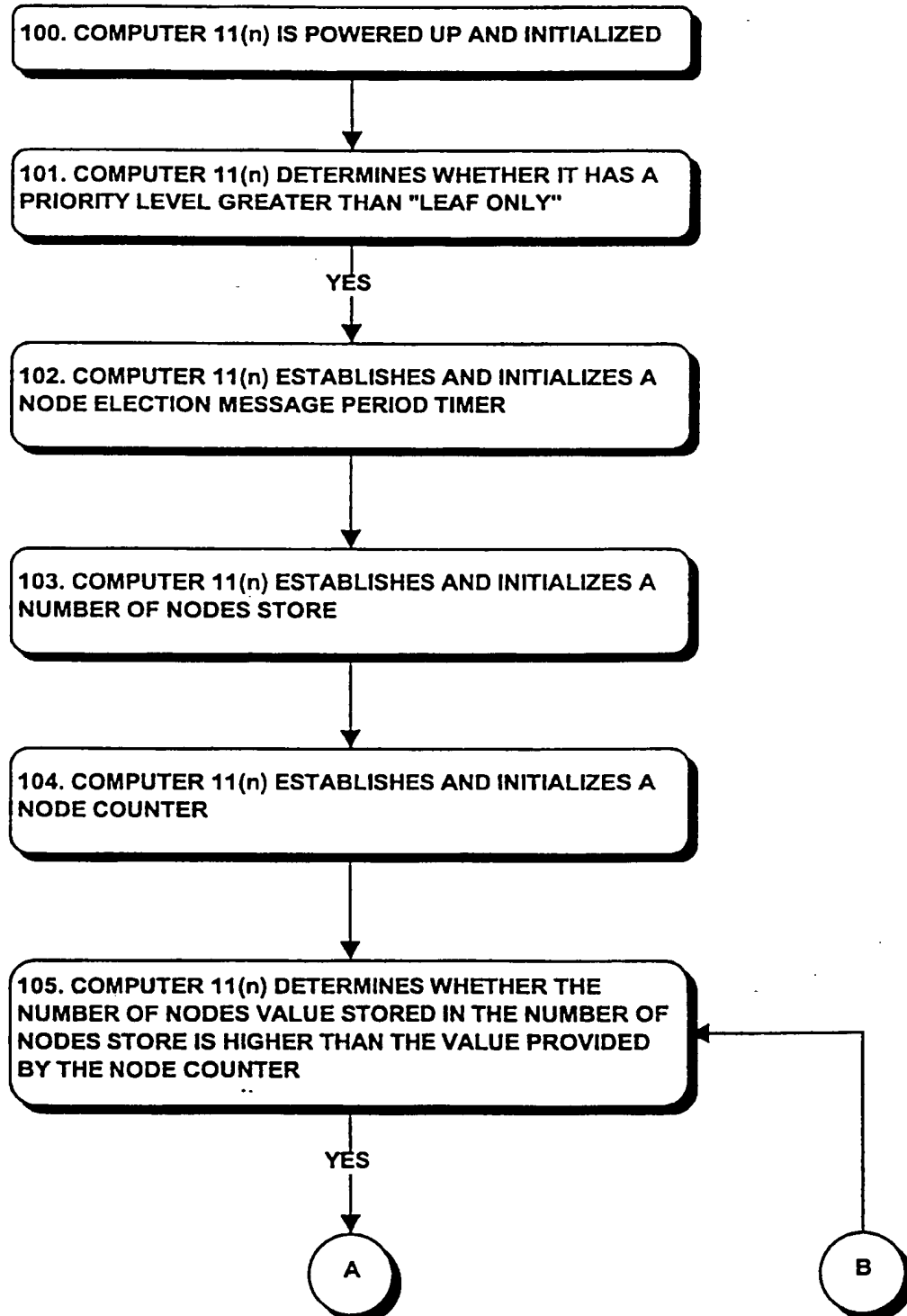


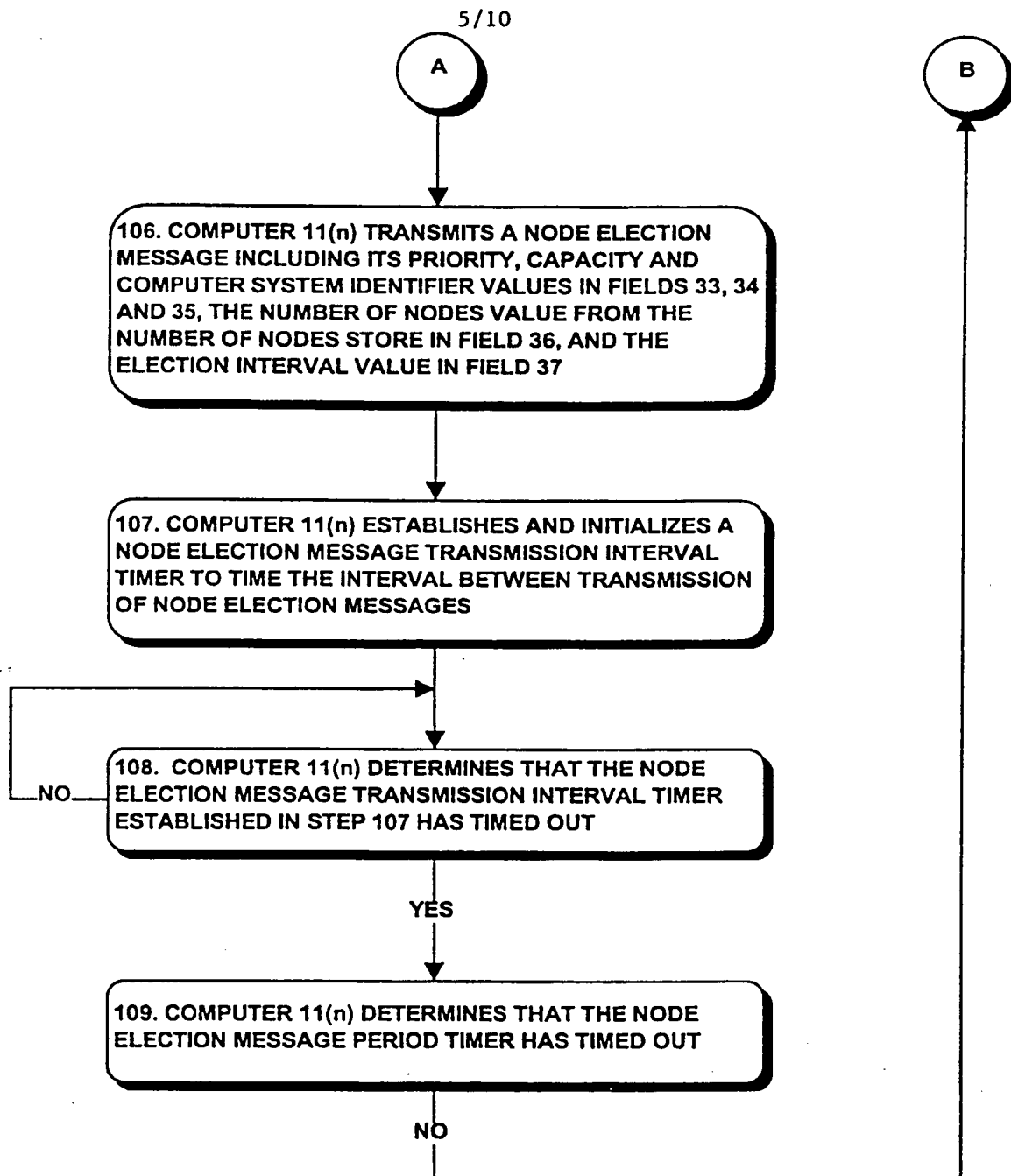
FIG. 1

FIG. 2

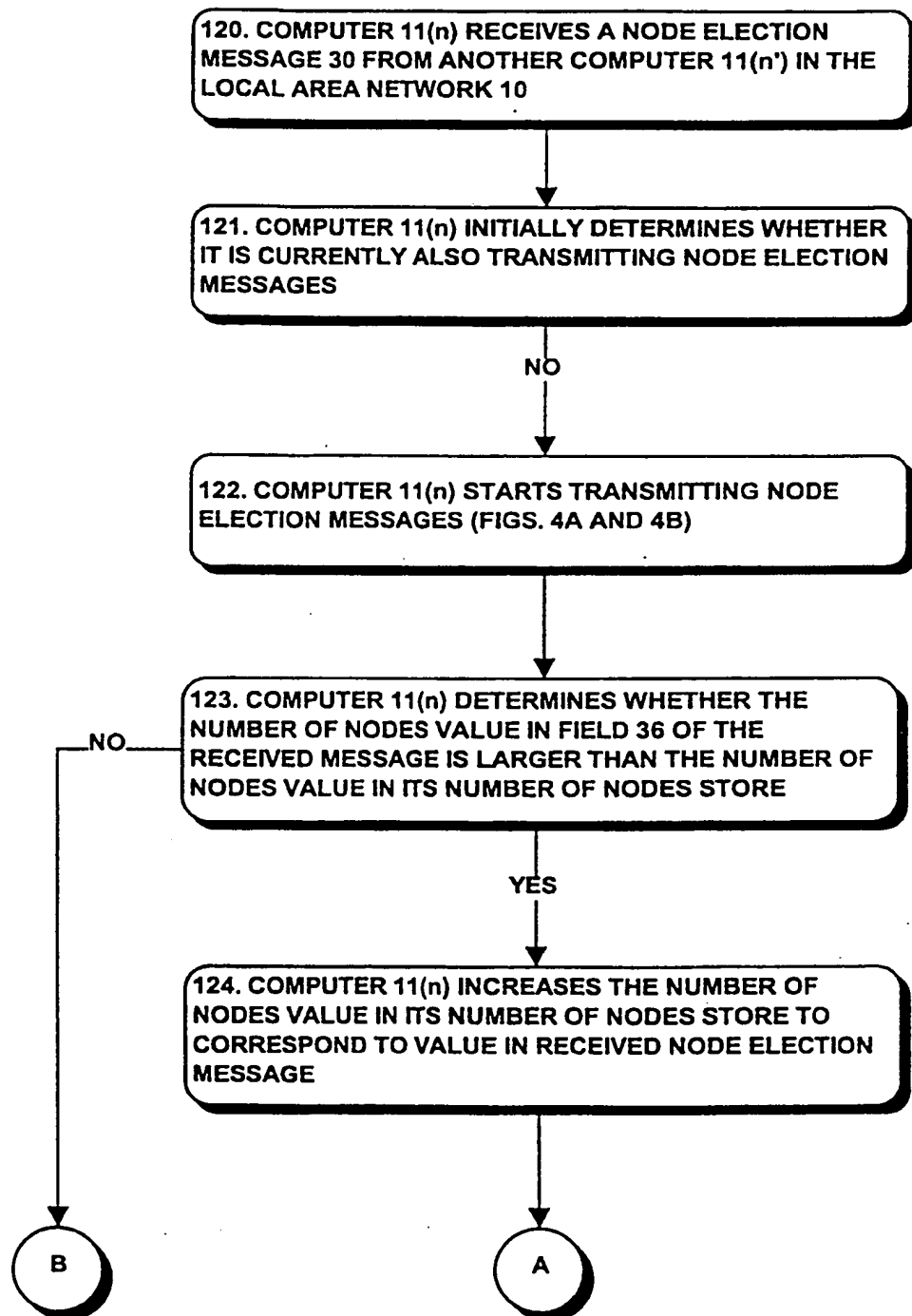


^{4/10}
FIG. 4A



FIG. 4B

6/10

FIG. 4C

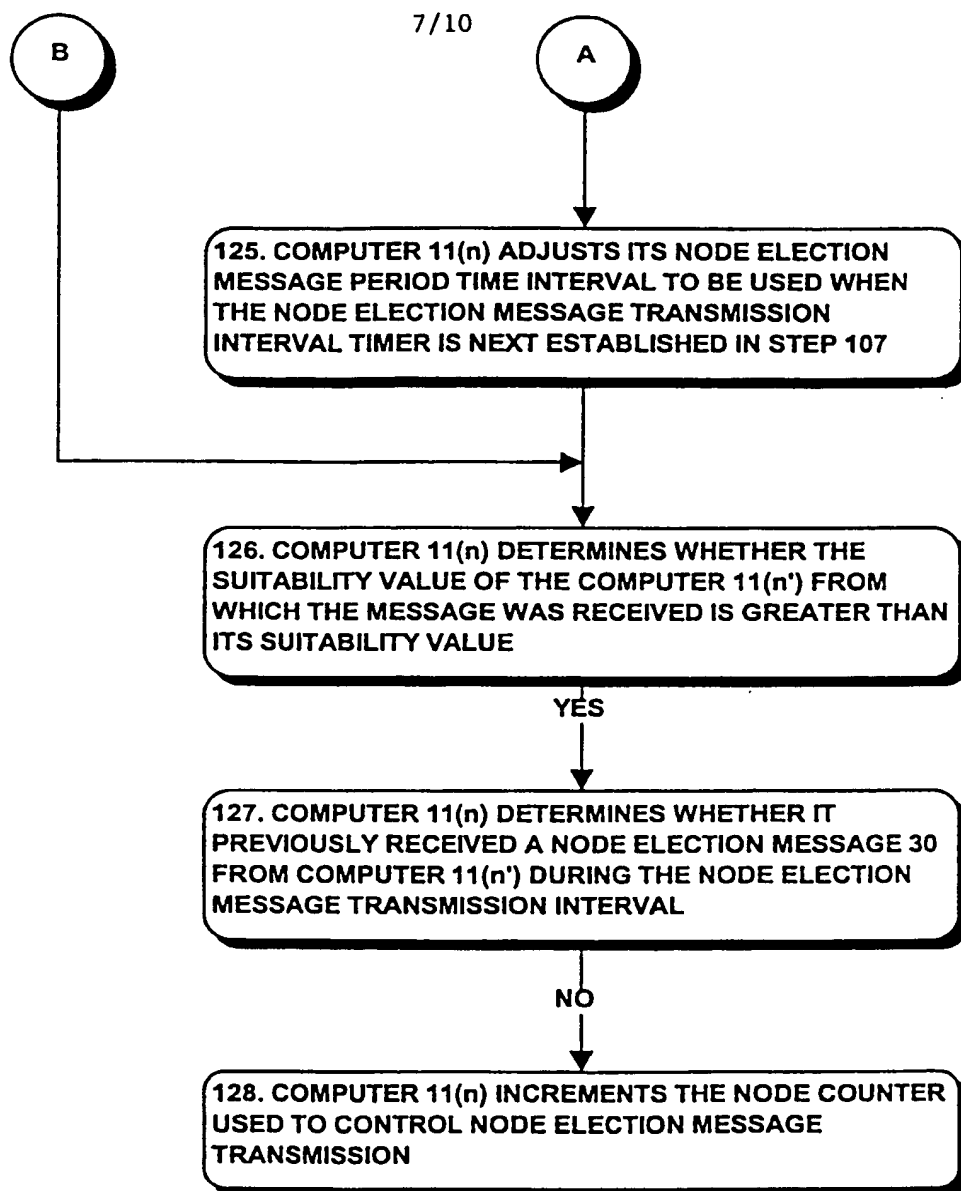
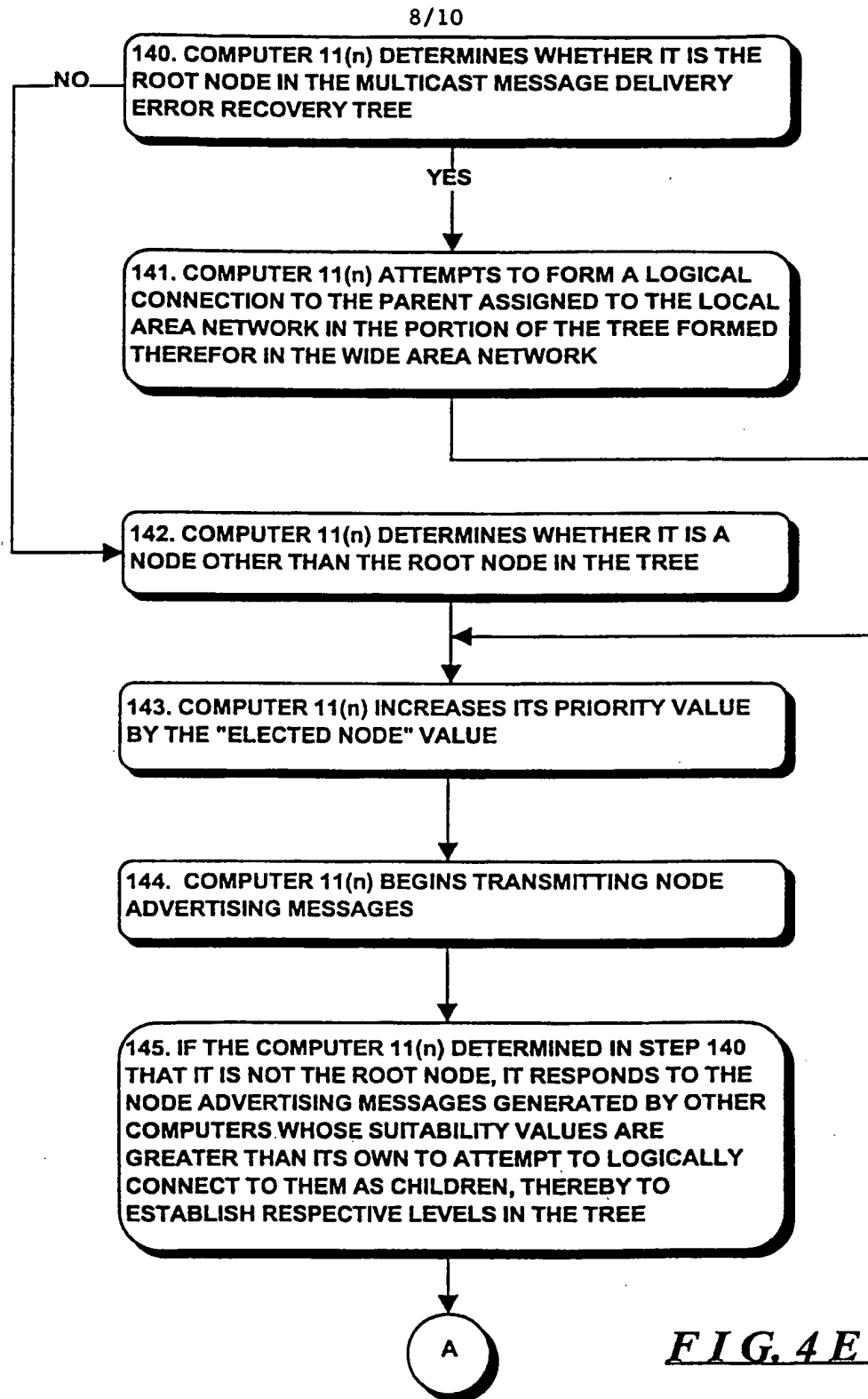
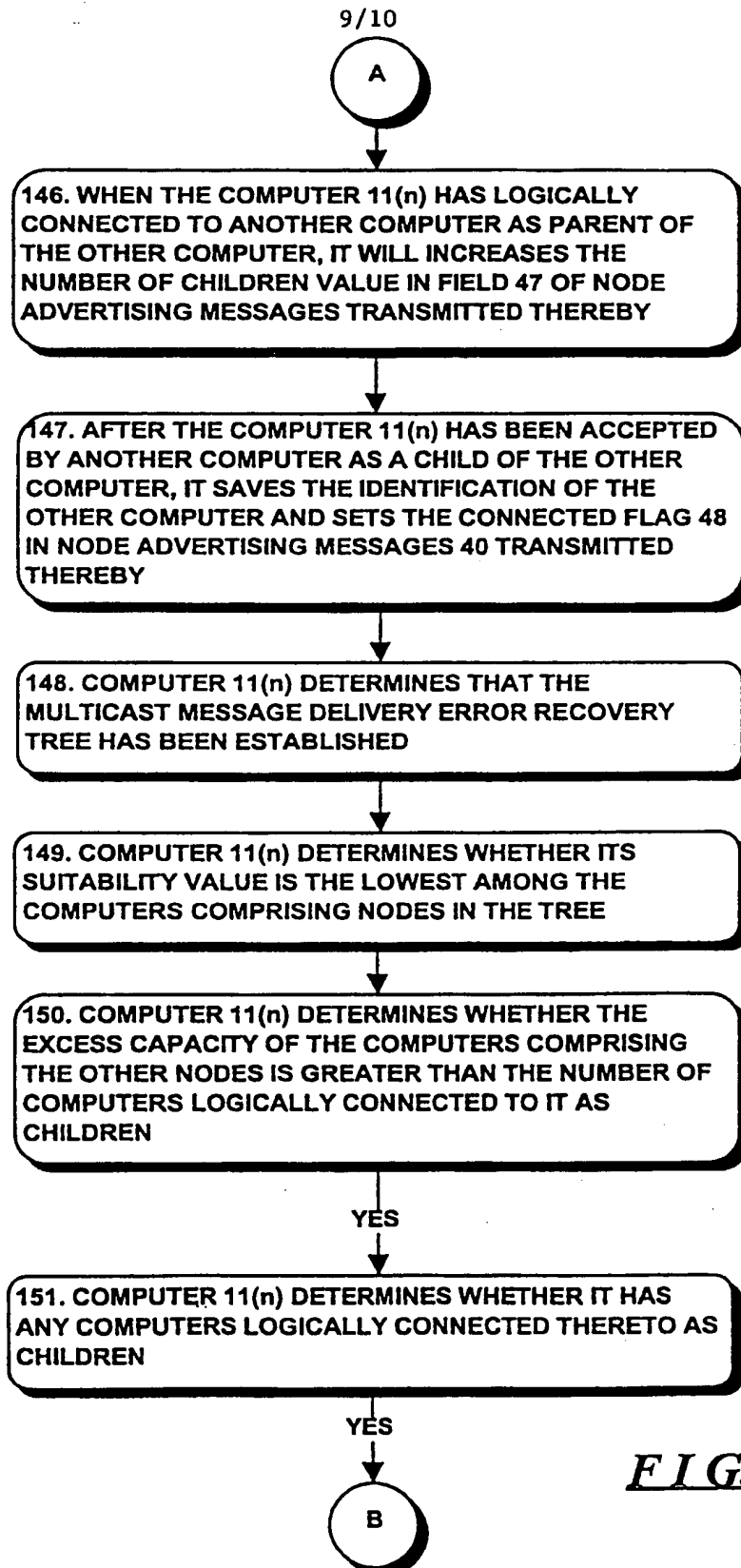
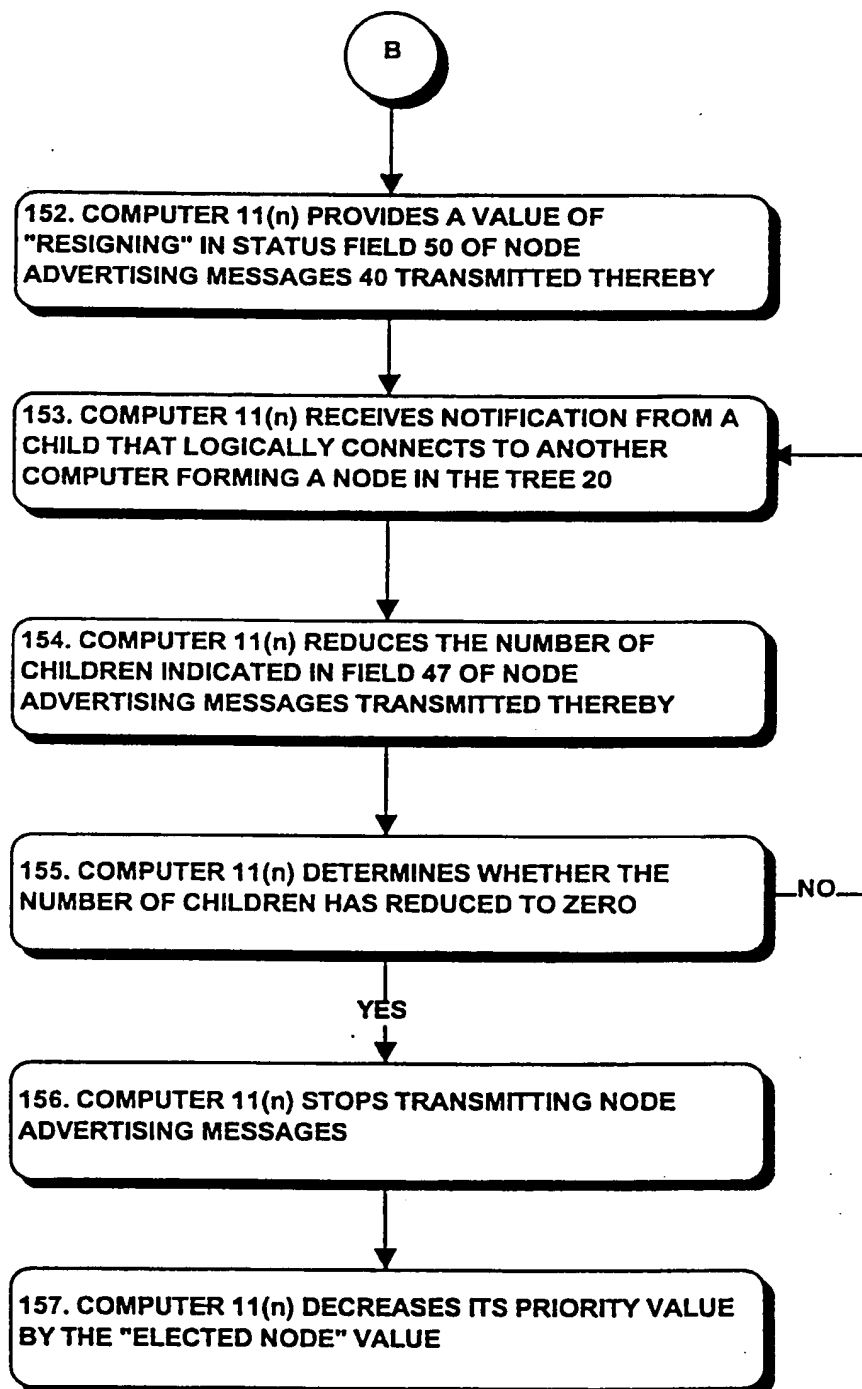


FIG. 4D

*FIG. 4E*

*FIG. 4F*

10/10

*FIG. 4G*

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 99/07750

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 H04L12/18

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	KIM J L ET AL: "A DISTRIBUTED ELECTION PROTOCOL FOR UNRELIABLE NETWORKS" JOURNAL OF PARALLEL AND DISTRIBUTED COMPUTING, vol. 35, no. 1, 25 May 1996 (1996-05-25), pages 35-42, XP000591142 ISSN: 0743-7315	1,2,8, 11-16, 22, 25-30, 36,39-42
A	page 35, left-hand column, line 1 - right-hand column, line 11 page 37, left-hand column, line 14 - page 38, right-hand column, line 12 --- -/--	3,9,10, 17,23, 24,31, 37,38

☒ Further documents are listed in the continuation of box C.

☐ Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

20 August 1999

Date of mailing of the international search report

06/09/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

RAMIREZ DE AREL..., F